



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 136079

TO: Bao-Qun Li  
Location: REM-3D24/3C18  
Art Unit: 1648  
Monday, November 01, 2004

Case Serial Number: 10/089292

From: Edward Hart  
Location: Biotech-Chem Library  
REM-1A55  
Phone: 571-272-2512

edward.hart@uspto.gov

### Search Notes

Examiner Li,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

1999.09/30

**This Page Blank (uspto)**

136079

STIC-Biotech/ChemLib

From: Li, Bao-Qun  
Sent: Tuesday, October 26, 2004 7:36 AM  
To: STIC-Biotech/ChemLib  
Subject: FW: 10,089,292

-----Original Message-----

From: Li, Bao-Qun  
Sent: Tuesday, October 26, 2004 7:31 AM  
To: STIC-Biotech/ChemLib  
Subject: 10,089,292

Please do the nucleic acid sequence homology and interference search against the amino acid sequence of SEQ ID NO:  
2. Thanks.  
Bao Qun Li  
Art unit 1648.  
Tel.20904.  
**REM, C18.**

\*\*\*\*\*

STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2- \_\_\_\_\_  
Date Searcher Picked up: \_\_\_\_\_  
Date Completed: \_\_\_\_\_  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search

NA Sequence: # \_\_\_\_\_  
AA Sequence: # \_\_\_\_\_  
Structure: # \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_

**This Page Blank (uspto)**

## Protein Sequence Searches - 10/8/04

All of the sequence databases on the ABSS have been updated. A change has occurred in the protein databases.

- Two protein databases, SPTREMBL and SwissProt, are now produced as a single, merged database called UniProt.
- Results from UniProt have the file extension **.rup**.
- Sequences in UniProt are identified by the same ID that had been used in SPTREMBL or SwissProt.
- In instances where the database curators have determined that an SPTREMBL record and a SwissProt record represent the same sequence, the two records have been merged into one. Both IDs are present in the record. Any differences found between the two sequences are recorded in the FT (feature table) fields.

If you have any questions regarding these changes or your results, please contact any STIC searcher.

Pending Nucleic Acid and Pending Amino Acid database searches generate two sets of results each. The Pending databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions **.rnpm** and **.rnpn**

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions **.rapm** and **.rapn**

***Because they contain data that is confidential, the results of Pending database searches should not be left in the case .***

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 28, 2004, 10:18:59 ; Search time 2034 Seconds  
(without alignments)  
3815.960 Million cell updates/sec

Title: US-10-089-292A-2  
Perfect score: 1095  
Sequence: 1 QLFVSRPVSANGEPVTKLY.....SLGAGPVSISAVAVLAPPPR 213

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-C=/cgn2.1/USPRO\_spool/US10089292/runat\_27102004\_153928\_10167/app\_query.fasta\_1.391  
-DB=EST -QFMT=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOPCL=0 -LOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10089292 @CGN 1.1 2607 @runat\_27102004\_153928\_10167 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEOURY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hcc:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gsl1:\*  
9: gb\_gsl2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	101	9.2	889	5 BQ681065	BQ681065 AGENCOURT
C 2	101	9.2	927	5 BUI74132	BUI74132 AGENCOURT
C 3	96.5	8.8	751	4 BM290684	BM290684 EST577218
C 4	96	8.8	935	5 BQ679041	BQ679041 AGENCOURT
C 5	95.5	8.7	665	6 CD901115	CD901115 G356.102N
C 6	95	8.7	791	6 CD575859	CD575859 UCPT01.0
C 7	94	8.6	609	4 BM291795	BM291795 EST574337
C 8	94	8.6	904	6 CD792824	CD792824 EST664185
C 9	93	8.5	797	4 BM291070	BM291070 EST577604

10	92.5	8.4	553	4	BM676609	UI-E-E10-
11	92.5	8.4	613	5	BQ684257	UI-CF-ENO
12	92.5	8.4	613	7	CN887504	010512AAW
13	92.5	8.4	695	6	CB851742	UI-CF-EN1
14	92.5	8.4	717	4	BM683536	UI-E-E01-
15	92.5	8.4	909	4	BM801700	AGENCOURT
16	92.5	8.4	1046	5	EX375570	EX375570
17	92.5	8.4	1091	4	BG167780	602339914
18	92	8.4	571	2	AW749973	QV3-BT053
19	92	8.4	688	1	AL821981	AL821981
20	92	8.4	1047	4	BG698844	602703249
21	92	8.4	9183	9	AY400947	Homo sapi
22	91.5	8.4	635	5	BUT73632	UI-B-CL1-
23	91.5	8.4	705	4	BG855589	1024043C0
24	91	8.3	827	7	CK136482	MM2_1_2_F
25	91	8.3	881	6	CD796381	EST667742
26	90.5	8.3	9183	9	AY400948	Pan trogl
27	90.5	8.3	687	4	BJ618104	BJ618104
28	90.5	8.3	700	7	CK863106	34318 In
29	90.5	8.3	731	9	CC615969	OGULV09TV
30	90	8.2	745	2	BF667907	602122145
31	90	8.2	977	6	CA787406	AGENCOURT
32	89.5	8.2	484	1	AL121209	DKF2p762C
33	89.5	8.2	564	7	CN889248	010604AAW
34	89.5	8.2	565	7	CN890116	010617AAW
35	89.5	8.2	577	4	BM714229	UI-E-EJ0-
36	89.5	8.2	590	7	CN947302	020802AVB
37	89.5	8.2	605	6	CB660993	OSJNED03E
38	89.5	8.2	619	7	CN908753	030122ABL
39	89.5	8.2	648	7	CN889030	010604AAW
40	89.5	8.2	649	7	CN908826	030122ABL
41	89.5	8.2	652	7	CN881070	010425AAS
42	89.5	8.2	776	7	CO389874	AGENCOURT
43	89	8.1	468	1	AU190012	AU190012
44	89	8.1	494	6	CD489573	T25_C02_T
45	89	8.1	520	4	BG963333	602827756

ALIGNMENTS

RESULT' 1  
BQ681065/c  
LOCUS BQ681065 889 bp mRNA linear EST 15-JUL-2002  
DEFINITION AGENCOURT 8208181 NIH\_MGC\_112 Homo sapiens cdna clone IMAGE:6262741  
5', mRNA sequence.  
ACCESSION BQ681065  
VERSION BQ681065.1 GI:21793744  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 889)  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: DCTD/DTF  
CDNA Library Preparation: Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LUCM2425 row: i column: 14  
High quality sequence stop: S75.  
Location/Qualifiers  
1. .889  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6262741"





[illegible]

```

Db 411 AGTGAGGACAGTCCACACCCGACCTGGACCCACCCACCCCTGGGTCTGTCCATCTCAGT 352
Qy 103 AspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeuAsp 122
Db 351 CCCGCGCTTGAGCCTCTGGGCCAAAGCCACCTCTCTGAGCAGGAGCGGAGCGAAGAC 292
Qy 123 TrpThrLysValThrLeuAspGlyArgProLeuSer-----ThrIleGlnGlnTyrSer 140
Db 291 TGGGAGCAGCAGCAGCGGAGCAGCAGCGGCCCATGAGCCACCTCCACCTCCAGATGGT 232
Qy 141 LysThrPhePheValLeuProLeu 148
Db 231 CAGAGTTACATGGTCACTCCCTG 208

RESULT 5
CD901115/c 665 bp mRNA linear EST 14-JUL-2003
LOCUS G356.102N03F010913 G356 Triticum aestivum cDNA clone G356102N03,
DEFINITION mRNA sequence.
ACCESSION CD901115
VERSION CD901115.1 GI:32675443
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 665)
REFERENCE Genoplate.
AUTHORS Genoplate, a major partnership french program in plant genomics
TITLE Unpublished (2003)
JOURNAL Contact: Genoplate
COMMENT Genoplate
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'genoplate' (http://www.genoplate.com)
and http://genoplate-info.infobiogen.fr.

FEATURES
source
1..665
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="G356102N03"
/tissue_type="grain (356 degrees per day after
pollination)"
/clone_lib="G356"

ORIGIN
Alignment Scores:
Pred. No.: 3..18 Length: 665
Score: 95.50 Matches: 53
Percent Similarity: 35.14% Conservative: 25
Best Local Similarity: 23.87% Mismatches: 65
Query Match: 79 Indels: 79
DB: 6 Gaps: 10

US-10-089-292a-2 (1-213) x CD901115 (1-665)

Qy 2 LeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyrThr 21
Db 625 CTAATTACAGT---CCGGTCTCAGC-----CCACGCTGTTTTTTGTTTCC 581
Qy 22 SerValGluAsnAlaGlnAspLysGlyIleAlaIleProHisAspLeuGly 41
Db 580 GGAACA----- 575
Qy 42 GluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrPro 61
Db 574 -----GATGACAGCATCGTCGCTTCCTCTCCTGCT 542

```

```

Qy 62 SerProLaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSer 81
Db 541 GCCCTCTCCCTCGCTCTCCC----- 521
Qy 82 LeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyrVal 101
Db 520 -----GAGCTCGCTCGGGGTCCAGCGCGCGGTGAGGTGATGTTG 482
Qy 102 SerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeu 121
Db 481 CTCCTCCACCATGAGCGCGGACGTGTCGTCGCGTGGAGGCGCGGCGTCCGGGCATG 422
Qy 122 AspTrpThrLysValThrLeu-----AspGlyArgProLeuSerThrIleGlnGln 138
Db 421 TCATGGACCTCGATGGGCGGAGCGCATCCCGTCGGGCGAGTACCTGCCACGCTGCAGATG 362
Qy 139 TyrSerLysThrPhePheValLeu-----Pro-LeuArgGln 150
Db 361 AAGTAGTAGACCTTGGCGCTCGCTGAGCAGAGTGGGATATCACGCGCGCGGAGGAGCGG 302
Qy 150 YLysLeuSerPheTrp---GluAlaGlyThrThrLysAlaGlyTyrPro-TyrAsnTyrA 169
Db 301 TAGATGGGTTTCTGGGGTCGCACATGTCGAACCCGCGCTGTTCACTCCACCATC- 243
Qy 169 snThrThrAlaSerAspGlnLeuValGluAsnAlaAlaGlyHisArgValAlaIleS 189
Db 242 -----CAGCTTCTTGTCTGTCGCGCGTGTAGAA---CACAGATA----- 210
Qy 189 erThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaValLeuA 209
Db 209 -----GTCTCTCTTGGAGAGGCGCGTGGCTCTTGGCCCGCAGTCCGAGTAGTTG 161
Qy 209 la 209
Db 160 CT 159

RESULT 6
CD575859 791 bp mRNA linear EST 12-JUN-2003
LOCUS UCRPT01_02bb01_g3 Poncirus trifoliata CTV-challenged cDNA library -
DEFINITION AGI Poncirus trifoliata cDNA clone UCRPT01_02bb01, mRNA sequence.
ACCESSION CD575859
VERSION CD575859.1 GI:31671761
KEYWORDS EST.
SOURCE Poncirus trifoliata
ORGANISM Poncirus trifoliata
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Sapindales; Rutaceae; Poncirus.
1 (bases 1 to 791)
Roose,M.L., Ye,X., Federici,C.F., Close,T.J., Fenton,R.D.,
Wanamaker,S., Collura,K., Feuerbacher,O., Kim,H.R., Kudrna,D.,
Wing,R. and Yu,Y.
Development of EST Resources and New Genetic Markers for California
Citrus - Poncirus trifoliata CTV-challenged phloem - AGI
Unpublished (2003)
Contact: Mikeal Roose
Department of Botany & Plant Sciences, University of California
Riverside, CA, 92521-0124, USA
Tel: 9097874137
Fax: 9097874437
Email: mikeal.roose@ucr.edu
Seq primer: T3.
Location/Qualifiers
1..791
/organism="Poncirus trifoliata"
/mol_type="mRNA"
/cultivar="Pomeroy OP"
/db_xref="taxon:37690"
/clone="UCRPT01_02bb01"
/tissue_type="Phloem"
/dev_stage="10 - 30 cm shoots"
/lab_host="E. coli TJC121"

FEATURES
source

```



```
QY 28 GlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValValIle 47
Db 123 GAAATCAGGTATGATGAGGCTTCTGTACAAATGACACCAAGAGTCGACCGCTGCTGTG 182
QY 48 GlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSerArg 67
Db 183 GCTAAAGTACGGCTCTTTGGCAGCAAGATGCCCTTACTGACCGCCCTGTACACCTAGA 242
QY 68 -----ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLeuThr 83
Db 243 GATGAAGTCTACGAGTACATATCTTCAGGCGCCATGACATCAAGACCTTCGTGTGCC 302
QY 84 AlaAlaGluTyrAspGlnSerThrTyr-----GlySerSerThrGlyProValTyrVal 101
Db 303 GAGCCTCCGAGACCCAGTCACCTCTGCGCGAGGCGCTCACAAATGATCCTGCCATTGT 362
QY 102 SerAspSerValTrpLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSerLeu 121
Db 363 CAACATTCAGCTACACCCATTGCTGCTGAGCAGGCTTTTCTGCGCGCTGCG----- 413
QY 122 AspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140
Db 414 ---TACACCAGGCACCTCGACCATATGCTCCAGTGGGTACTCTGCCGTCAATTTCA 467

RESULT 8
CD792824
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

CD792824 904 bp mRNA linear EST 01-JUL-2004
EST664185 RAB Rhipicephalus appendiculatus cDNA clone RAB659 5'
end. mRNA sequence.
CD792824
CD792824.1 GI:49548498
EST.
Rhipicephalus appendiculatus
Rhipicephalus appendiculatus
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
Parasitiformes; Ixodidae; Ixodidae; Rhipicephalus.
Nene, V., Lee, Y., Skilton, R., Mwaura, S., Quackenbush, J., Gardner, M.
and Bishop, R.
1 (bases 1 to 904)
An index of genes transcribed in the salivary glands of
Rhipicephalus appendiculatus
Unpublished (2003)
Other ESTs: EST664184
Contact: Vishvanath Nene
Parasite Genomics
The Institute for Genomic Research
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-610-5968
Fax: 301-838-0208
Email: nene@tigr.org
Seq primer: M13 reverse.

FEATURES
source
1. 904
/organism="Rhipicephalus appendiculatus"
/mol_type="mRNA"
/strain="Muguga"
/db_xref="taxon:34631"
/clone="RAB659"
/dev_stage="Adult"
/lab_host="E. coli strain DH10B-Tona"
/clone_lib="RAB"
/notes="Organ: Salivary glands; Vector: pCMVSPORT6.0.ccdB;
Salivary glands were dissected on day four after
initiation of feeding. Total RNA was prepared using acid
guanidium thiocyanate-phenol-chloroform extraction. The
cDNA library was custom prepared by Invitrogen
Corporation. Briefly, first strand cDNA was primed using
oligo(dT) containing a NotI site. Size fractionated double
stranded cDNA was ligated to EcoRV-NotI cleaved vector and
electroporated into E.coli. Library RAB was made from
ticks infected with Theileria parva."
```

```
Alignment Scores:
Pred. No.: 7.15 Length: 904
Score: 94.00 Matches: 48
Percent Similarity: 43.17% Conservative: 31
Best Local Similarity: 26.23% Mismatches: 73
Query Match: 8.58% Indels: 32
DB: 6 Gaps: 8

US-10-089-292A-2 (1-213) x CD792824 (1-904)
QY 7 ProValValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerVal-GluAsnAl 26
Db 83 CCGTACTTGGAGCGCAAAATAAGCCTAATATCAAGATCAGAGATTCGGTATGAG----- 137
QY 26 aGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeuGlyGluSerArgValVa 46
Db 138 -----GGCATTTCTGTACACATTTGATACCAAGAGTCAACCGTGGC 178
QY 46 IleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerProAlaProSe 66
Db 179 TTTGGCTAAAGTGGCGTCTTTGGCAGCAAGATCGCCCACTGATCGCGCAGTGGCACC 238
QY 66 rArg-----ProPheSerValLeuArgAlaAsnAspValLeuTrpLeuSerLe 82
Db 239 TAGAGATGAAGTTTACGAGTACATCTCTCAGGCGCCATGACATCAAGATCTCAGAGT 298
QY 82 uThrAlaAlaGluTyrAspGlnSerThrTyr-----GlySerSerThrGlyProValTy 100
Db 299 GTCGAGCGCTCTTAAACCGCAGTCAACATTACCTGGTGGCCTTAACAAATGATCTCTAT 358
QY 100 rValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSe 120
Db 359 TGTTCAACATTCAGCTACT-----CCTGTTGGCACTGGAGCGCTTTTCTGCAACGCGC 412
QY 120 rLeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnTyrSe 140
Db 413 A-----TATGACCAACCGCAGCTTCATATGTCCTGATGGGAGTTACCGTCTTATTC 466
QY 140 rLysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrTh 160
Db 467 TATGCTCC-TCAATATAC-----CTGCAGCGCTTCACAGCAGCAGCCACC 510
QY 160 rLysAla-----GlyTyrPro-----TyrAsnTyrAsnThrThrAlaSe 173
Db 511 CAAGTCTAGTGGATCCCAACAATCGGTTCCCACTCTGGCAGTACAACTCCTGCACAGC 570
QY 173 rAspGln 175
Db 571 GAAGAG 577

RESULT 9
BM291070 797 bp mRNA linear EST 01-JUL-2002
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
COMMENT

BM291070
EST577604 AVSG Amblyomma variegatum cDNA clone AVAAT96 5' end, mRNA
sequence.
BM291070
BM291070.1 GI:21641037
EST.
Amblyomma variegatum
Amblyomma variegatum
Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;
Parasitiformes; Ixodidae; Ixodidae; Amblyomma.
1 (bases 1 to 797)
Nene, V., Lee, D., Quackenbush, J., Skilton, R., Mwaura, S.,
Gardner, M.J. and Bishop, R.
AvGT, an index of genes transcribed in the salivary glands of the
Ixodid tick Amblyomma variegatum
Int. J. Parasitol. 32 (12), 1447-1456 (2002)
2281296
12392910
Contact: Vish Nene
Parasite Genomics Group
The Institute for Genomic Research
```

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-610-5968

Fax: 301-838-0208

Email: nene@tigr.org

Seq primer: M13 reverse.

Location/Qualifiers

# FEATURES

source

1..797

/organism="Amblyomma variegatum"

/mol\_type="mRNA"

/db\_xref="taxon:34610"

/clones="AVAAAT96"

/tissue\_type="Salivary glands"

/dev\_stage="Adult"

/lab\_host="E.coli strain DH10B-Tona"

/clone\_lib="AVSG"

/note="Vector: pCMV-SPORT6.1; Salivary glands were

dissected on day five after initiation of feeding. Total

RNA was prepared using acid guanidinium

thiocyanate-phenol-chloroform extraction. The cDNA library

was custom prepared by Invitrogen Corporation. Briefly,

first strand cDNA was primed using oligo(dT) containing a

NotI site. Size fractionated double stranded cDNA was

ligated to EcoRV-NotI cleaved vector and electroporated

into E.coli."

## ORIGIN

### Alignment Scores:

Pred. No.: 7.82 Length: 797  
Score: 93.00 Matches: 56  
Percent Similarity: 37.04% Conservative: 34  
Best Local Similarity: 23.05% Mismatches: 105  
Query Match: 8.49% Indels: 49  
DB: 4 Gaps: 9

US-10-089-292A-2 (1-213) x BM291070 (1-797)

Qy 4 TyrSerArgProValSerAlaAsnGlyGluProThrValLysLeuTyrThrSerVal 23  
Db 82 TACCGGTATACCGTACTTGGCGAGCAAAATAGCCTATATCAAGTCTGAATCAGGTA 141  
Qy 24 -GluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspLeuGlyGluSe 43  
Db 142 TGAA-----GGCATCTGTACACAAATGTACCAAGATC 177  
Qy 43 rArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThrProSerPr 63  
Db 178 GACCGTTGCTCGGTAAAGTACGCTCTTGGCACAGAGATCGCCCTACTGACCGGCC 237  
Qy 63 oAlaProSerArg-----ProPheSerValLeuArgAlaAsnAspValLeuTr 79  
Db 238 TGTAGCACCTAGAGATGAAGTCTACGAGTACATTATCTTCAGGGCCCATGACATCAAGA 297  
Qy 79 pLeuSerLeuThrAlaAlaGluTyrAspGlnSerThrTyr-----GlySerSerThrGl 97  
Db 298 CTTTGTGTGTCGAGCGCTCCGAGCCCGAGTCAACTCTGCCGAGGAGCCCTCACAATGA 357  
Qy 97 yProValTyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaVa 117  
Db 358 TCCTGCCATTGTTCAACATTACAGTACACCCATTGGTCTGGAGCAGGCTTCTTGCGCC 417  
Qy 117 lAlaArgSerLeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIle-- 136  
Db 418 TGG-----TACAACCGAGCACCTGCACCATATGTCAGTGGGTACTCTGCC 465  
Qy 137 -----GlnGlnTyrSerLysThrPhePheVal----- 145  
Db 466 GTCATATTCAATGGCTCCACAGTACTCTTC-ACAACCTTCACAGCCACCGACGAGCAGCC 524  
Qy 146 -----LeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrLysAlaGl 163  
Db 525 TAGTGGGTCCCGCAATC---AGTTCTCAGTCTGGAAGCACAACTCCACGACGCGGA 581  
Qy 163 yTyrProTyrAsnTyr-----AsnThrThrAl 172

Db 582 AGACCCCAACAATAGATGCTGGAGTACAGGTTTCACTCCCTCCCACTGAAAACACAGT 641  
Qy 172 aserAspGlnLeuLeuValGluAsnAlaAlaGlyHisArgVal-----AlaIleSe 189  
Db 642 CTCCTCTGTGGTGGCTCTCAGCGCAGCCATGGGGCAGCAAGTTCTGCAGGACCTAGCAGG 701  
Qy 189 rThrTyrThrThrSerLeuGlyAlaGlyProValSerIleSerAlaValAlaValLeuAl 209  
Db 702 CAGCATGTCAGCAACATGGGCAGCAGCATCACCAGCAGCATAGCCACAGCAGCACCAG 761  
Qy 209 aProPro 211  
Db 762 CAGACCC 768

## RESULT 10

BM676609

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

BM676609 553 bp mRNA linear EST 27-FEB-2002  
UI-E-EJ0-ahq-h-20-0-UI.s2 UI-E-EJ0 Homo sapiens cDNA clone  
UI-E-EJ0-ahq-h-20-0-UI 3', mRNA sequence.

BM676609

BM676609.1 GI:18986505

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 553)

Bonaldo,M.F., Lennon,G. and Soares,M.B.

Normalization and subtraction: two approaches to facilitate gene

discovery

Genome Res. 6 (9), 791-806 (1996)

97044477

8889548

Contact: Soares, MB

Coordinated Laboratory for Computational Genomics

University of Iowa

375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA

Tel: 319 335 8250

Fax: 319 335 9565

Email: bento-soares@uiowa.edu

Tissue Procurement: Dr. Gregg Hageman

cDNA Library preparation: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Researchers may obtain clones from Research

Genetics (www.resgen.com).

Seq primer: M13 Forward

PolyA=Yes.

Location/Qualifiers

1..553

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clones="UI-E-EJ0-ahq-h-20-0-UI"

/tissue\_type="fetal eyes, lens, eye anterior segment,

optic nerve, retina, Retina Foveal and Macular, RPE and

Choroid"

/dev\_stage="fetal and adult"

/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"

/clone\_lib="UI-E-EJ0"

/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a

modified polylinker; Site 1: EcoR I; Site 2: Not I;

UI-E-EJ0 is a subtracted cDNA library constructed

according to Bonaldo, Lennon and Soares, Genome Research,

6:791-806, 1996. First strand cDNA synthesis was primed

with an oligo-dT primer containing a Not I site. Double

stranded cDNA was ligated to an EcoR I adaptor, digested

with Not I, and cloned directionally into pT7T3-Pac

vector. The oligonucleotide used to prime the synthesis of

first-strand cDNA contains a library tag sequence that is

located between the Not I site and the (dT)18 tail. The

sequence tags for this library are: fetal eyes,

AGAAATCAAGA; lens, CGATTACGGA; eye anterior segment, AATGCCGAT; optic nerve, CCATTAGTG; retina, CCAGC; Retina Foveal and Macular, GTC; RPE and Choroid, ACCTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).  
TAG\_TISSUE=human optic nerve  
TAG\_LIB=UI-E-EJO  
TAG\_SEQ=CCATTAGTG

## ORIGIN

Alignment Scores:  
Pred. No.: 5.41 Length: 553  
Score: 92.50 Matches: 30  
Percent Similarity: 43.64% Conservativity: 18  
Best Local Similarity: 27.27% Mismatches: 41  
Query Match: 8.45% Indels: 21  
DB: 4 Gaps: 4

US-10-089-292A-2 (1-213) x BM676609 (1-553)

QY 38 lleAspLeuGlyGluSerArgValValIleGlnAspTyrAspAsnGlnHis----- 54  
DB 204 CTGGATTGGGAGGTGGACAGGGGATCTTCAACCCCTACCATCTATCCAGGGTCTC 263  
QY 55 -----GluGlnAspArgProThrProSerProAlaProSerArgProPheSerValLeu 72  
DB 264 ATGTGTCAGCTGTCCAGACAGGTCCCAATCTGCTCCCATCTACCTCCCTGGGACATCTG 323  
QY 73 ---ArgAlaAsnAspValLeuThrPleuSerLeuThrAlaAlaGluTyrAspGlnSerThr 91  
DB 324 TACAGGATTCCTCAGTCTCTCTGG---GAACCTAACCCACCTGCTAGCCATCTAGCTTC 380  
QY 92 TyrGlySerSer-----ThrGly 97  
DB 381 TGGGCTCCACTGACGTGACGACAGTCCACACCCGACCTGGACCCCACTGGG 440  
QY 98 ProValTyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaVal 117  
DB 441 TCTGTCCATCTCAGTCCCGGCTGAGCTCTGGGCAAGCCACCTCTTCTGAGCAGGCA 500  
QY 118 AlaArgSerLeuAspTrpThrIysValThr 127  
DB 501 GGCAGAGCGAGAGACTGGGAGCAGCAGACA 530

## RESULT 11

BU684257  
LOCUS BU684257 613 bp mRNA linear EST 07-OCT-2002  
DEFINITION UI-CF-ENO-aco-e-22-0-UI-s1 UI-CF-ENO Homo sapiens cDNA clone  
ACCESSION UI-CF-ENO-aco-e-22-0-UI 3', mRNA sequence.  
VERSION BU684257  
KEYWORDS EST.  
SOURCE Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 613)  
Bonaldo, M.F., Lennon, G. and Soares, M.B.

## AUTHORS

Normalization and subtraction: two approaches to facilitate gene discovery

## TITLE

Genome Res. 6 (9), 791-806 (1996)

## JOURNAL

97044477

## MEDLINE

8889548

## PUBMED

Contact: McCray, PB

University of Iowa

2024 University of Iowa Med Labs, Iowa City, IA 52242, USA

Tel: 319 356 4866

Fax: 319 356 7171

Email: paul-mccray@uiowa.edu

Tissue Procurement: Dr. M. J. Welsh, University of Iowa

cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa

cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa

DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com) or from Open Biosystems (www.openbiosystems.com).  
Seq primer: M13 FORWARD  
POLYA-res.

## FEATURES

## source

Location/Qualifiers  
1..613  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-CF-ENO-aco-e-22-0-UI"  
/tissue\_type="Primary Lung Cystic Fibrosis Epithelial Cells"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/clone\_lib="UI-CF-ENO"  
/note="Organ: Lung; Vector: p7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-CF-ENO is a cDNA library containing the following tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into p7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (d)18 tail. The sequence tag for this library is CTGCTCAGGT.  
TAG\_TISSUE=Human Lung Epithelial Cell Lines untreated LPS  
6hr to LPS 24h  
TAG\_LIB=UI-CF-ENO  
TAG\_SEQ=CTGCTCAGGT"

## ORIGIN

## Alignment Scores:

Pred. No.: 6.23 Length: 613  
Score: 92.50 Matches: 30  
Percent Similarity: 43.64% Conservativity: 18  
Best Local Similarity: 27.27% Mismatches: 41  
Query Match: 8.45% Indels: 21  
DB: 5 Gaps: 4

US-10-089-292A-2 (1-213) x BU684257 (1-613)

QY 38 lleAspLeuGlyGluSerArgValValIleGlnAspTyrAspAsnGlnHis----- 54  
DB 203 CTGGATTGGGAGGTGGACAGGGGATCTTCAACCCCTACCATCTATCCAGGGTCTC 262  
QY 55 -----GluGlnAspArgProThrProSerProAlaProSerArgProPheSerValLeu 72  
DB 263 ATGTGTCAGCTGTCCAGACAGGTCCCAATCTGCTCCCATCTATCCCTGGGACATCTG 322  
QY 73 ---ArgAlaAsnAspValLeuThrPleuSerLeuThrAlaAlaGluTyrAspGlnSerThr 91  
DB 323 TACAGGATTCCTCAGTCTCTCTGG---GAACCTAACCCACCTGCTAGCCATCTAGCTTC 379  
QY 92 TyrGlySerSer-----ThrGly 97  
DB 380 TGGGCTCCACTGACGTGACGACAGTCCACACCCGACCTGGACCCCACTGGG 439  
QY 98 ProValTyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaVal 117  
DB 440 TCTGTCCATCTCAGTCCCGGCTGAGCTCTGGGCAAGCCACCTCTTCTGAGCAGGCA 499  
QY 118 AlaArgSerLeuAspTrpThrIysValThr 127  
DB 500 TGCAGAGCGAGAGACTGGGAGCAGCAGACA 529

## RESULT 12

CN887504

```

LOCUS               613 bp      mRNA      EST 04-JUN-2004
DEFINITION          010512AAWA002206HT (AAWA) Royal Gala 59 DAFB seeds Malus x
                    domesticA cDNA clone AAWA002206, mRNA sequence.
ACCESSION            CN887504
VERSION              CN887504
KEYWORDS             
SOURCE                CN887504.1 GI:48273746
                    EST.
                    Malus x domestica (cultivated apple)
                    Malus x domestica
                    Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
                    Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                    rosids; eurosids I; Rosales; Rosaceae; Maloideae; Malus.
                    1 (bases 1 to 613)
REFERENCE             Buning,L., Bowen,J., Crowhurst,R., Gleave,A., Jansen,B.,
                    McArtney,S., Newcomb,R., Ross,G., Snowden,K., Walton,E. and Yauk,Y.
                    HortResearch Apple EST Project
                    Unpublished (2004)
COMMENT               Contact: Gleave,A.
                    Sequencing Facility
                    The Horticulture and Food Research Institute of New Zealand Ltd
                    120 Mt Albert Rd, Mt Albert, Auckland, New Zealand
                    Tel: 00 64 09 815 4200
                    Fax: 00 64 09 815 4201
                    Email: est@hortresearch.co.nz.
                    Location/Qualifiers
                        1..613
                        /organism="Malus x domestica"
                        /mol_type="mRNA"
                        /db_xref="taxon:3750"
                        /clone="AAWA002206"
                        /tissue_type="Seed"
                        /dev_stage="59 days after full bloom"
                        /clone_lib=" (AAWA) Royal Gala 59 DAFB seeds"
                        /note="vector: pBluescript SK(-); Library sequenced by
                        Genesis Research & Development"
FEATURES              source

```

## ORIGIN

<b>Alignment Scores:</b>					
Pred. No.:	6.23	Length:	613		
Score:	92.50	Matches:	42		
Percent Similarity:	37.50%	Conservative:	21		
Best Local Similarity:	25.00%	Mismatches:	54		
Query Match:	8.45%	Indels:	51		
DB:	7	Gaps:	6		
US-10-089-292A-2 (1-213) x CN887504 (1-613)					
Qy	73	ArgAlaAsnAspValLeuThrPLeuSerLeuThrAlaAlaGlu-----	86		
Dd	10	AAGAAAAAGGAICTAGTCTGGGTTCGCCACAGGCCGCAGATGTTGTCGAACTTTTATC	69		
Qy	87	-----TyrAsp	88		
Dd	70	TATTTAGTGNAACCTGGCATGTTGTGCAGCTCTCTCAACAATAACCNTGGTGTAGAT	129		
Qy	89	GlnSerThrTyGlySerSerThrGlyProValTyrValSerAspSerValThrLeuVal	108		
Dd	130	GATTCAAACCTTATCCATCAACA-	153		
Qy	109	AsnValAlaThrGlyAlaGlAlaValAlaIahArgSerLeuAspTrpThrLysvalThrLeu	128		
Dd	154	GACTTAAGGACAGGA-----CGCTCTTTTAGATGGGCTAAAACTTGTTATG	198		
Qy	129	AspGlyArgProLeuSerThrIleGlnTyrSerLysThrPhePheValLeuProLeu	148		
Dd	199	GAGGGT-----GCTTCGATACCTCATTGTGTAAATGGAACAAAACCTCTTGATACCCCTTA	252		
Qy	149	ArgGlyLysLeuSerPheTrpGluAlaGlyThrThrLysAlaGlyTy-ProTyrAsnTyr	168		
Dd	253	ACGGGGCAATTAGTCCAGAGGATAGCTGTAAACAGGAGCTGGTTCAAGTCTTCATGCT	312		
Qy	169	AsnThrThrAlaSerAspGlnLeuLeu-----ValGluAsnAlaLaGlyHisArg---	185		
Dd	313	CAAGATACATCTAGCGTTCCAATTATGTTATGATTGTTGAAGAAGCTGGAAGGAACTTGAC	372		

```

Score: 92.50 Matches: 30
Percent Similarity: 43.64% Conservative: 18
Best Local Similarity: 27.27% Mismatches: 41
Query Match: 8.45% Indels: 21
DB: 6 Gaps: 4

US-10-089-292A-2 (1-213) x CB851742 (1-695)

QY 38 lleAspLeuGlyGluSerArgValValIleGlnAspTyrAspAsnGlnHis----- 54
Db 204 CTGGATTGGGAGGTGGAGAGGGGATCTTCAACCCCTACCATCTATCCAGGGCTC 263

QY 55 -----GluGlnAspArgProThrProSerProAlaProSerArgProPheSerValLeu 72
Db 264 ATGTGTCCAGCTGTCCAGAGCAGTCCCAATCTGCTCCCATCTACCTCGGACATCTCTG 323

QY 73 ---ArgAlaAsnAspValLeuThrLeuSerLeuThrAlaAlaGluTyrAspGlnSerThr 91
Db 324 TACAGGATTCCCTCAGTCTCTCTGG---GAACCTAACCCACCTGCTACCCATCTAGCTTC 380

QY 92 TyrGlySerSer-----ThrGly 97
Db 381 TGGGGCTCCACTGACAGCAGTGACAGCAGTCCACACCACTGGACCCCACTGGG 440

QY 98 ProValTyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaVal 117
Db 441 TCTGTCCATCTCAGTCCCGGCTGAGCTCTGGGCAAGCCACCTCTCTCTGAGCAGGCA 500

RESULT 14
BM683536 717 bp mRNA linear EST 27-FEB-2002
LOCUS UI-E-EO1-ajb-h-05-0-UI.s1 UI-E-EO1 Homo sapiens cDNA clone
DEFINITION UI-E-EO1-ajb-h-05-0-UI.3', mRNA sequence.
VERSION BM683536
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 717)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
Tissue Procurement: Dr. Gregg Hageman
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
Seq primer: M13 Forward
POLYA=Yes.

FEATURES
source
1. 717
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-E-EO1-ajb-h-05-0-UI"
/tissue_type="fetal eye"

/dev stages="fetal"
/lab host="DH10B (Life Technologies) (T1 phage resistant)"
/clone.lib="UI-E-EO1"
/notes="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-EO1 is a normalized cDNA library containing the
following tissue(s): fetal eye. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pT7T3-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
CGCGTATACC. This library was created for the program, Gene
Discovery in the Visual System, supported by National Eye
Institute (NEI).
TAG TISSUE=human fetal eye
TAG LIB=UI-E-EO1
TAG_SEQ=CGCGTATACC"

ORIGIN
Alignment Scores:
Pred. No.: 7.71 Length: 717
Score: 92.50 Matches: 30
Percent Similarity: 43.64% Conservative: 18
Best Local Similarity: 27.27% Mismatches: 41
Query Match: 8.45% Indels: 21
DB: 6 Gaps: 4

US-10-089-292A-2 (1-213) x BM683536 (1-717)

QY 38 lleAspLeuGlyGluSerArgValValIleGlnAspTyrAspAsnGlnHis----- 54
Db 204 CTGGATTGGGAGGTGGAGAGGGGATCTTCAACCCCTACCATCTATCCAGGGCTC 263

QY 55 -----GluGlnAspArgProThrProSerProAlaProSerArgProPheSerValLeu 72
Db 264 ATGTGTCCAGCTGTCCAGAGCAGTCCCAATCTGCTCCCATCTACCTCGGACATCTCTG 323

QY 73 ---ArgAlaAsnAspValLeuThrLeuSerLeuThrAlaAlaGluTyrAspGlnSerThr 91
Db 324 TACAGGATTCCCTCAGTCTCTCTGG---GAACCTAACCCACCTGCTACCCATCTAGCTTC 380

QY 92 TyrGlySerSer-----ThrGly 97
Db 381 TGGGGCTCCACTGACAGCAGTGACAGCAGTCCACACCACTGGACCCCACTGGG 440

QY 98 ProValTyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaVal 117
Db 441 TCTGTCCATCTCAGTCCCGGCTGAGCTCTGGGCAAGCCACCTCTCTCTGAGCAGGCA 500

RESULT 15
BM801700 909 bp mRNA linear EST 05-MAR-2002
LOCUS AGENCOURT 6458620 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5573496
DEFINITION 5', mRNA sequence.
ACCESSION BM801700
VERSION BM801700.1 GI:19118523
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 909)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

```



JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: LIA12320 row: k column: 01  
High quality sequence stop: 631.  
Location/Qualifiers  
1. .909  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5573496"  
/tissue\_type="duodenal adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 88"  
/note="Organ: small intestine; Vector: pCMV-SPORT6;  
Site 1: NotI; Site 2: SalI; Cloned unidirectionally;  
oligo-dT primed. Average insert size 1.767 kb. Library  
enriched for full-length clones and constructed by Life  
Technologies. Note: this is a NIH\_MGC Library."

FEATURES  
source

## ORIGIN

Alignment Scores:	10.7	Length:	909
Pred. No.:	92.50	Matches:	30
Score:	43.64%	Conservative:	18
Best Local Similarity:	27.27%	Mismatches:	41
Query Match:	8.45%	Indels:	21
DB:	4	Gaps:	4

US-10-089-292A-2 (1-213) x BM801700 (1-909)

Qy	38	IleAspLeuGlyGluSerArgValIleGlnAspTyrAspAsnGlnHis-----	54
Db	435	CTGGATTGGGAGGTGGACAGGGGATCTTCAACCCCTACCATCTCCTCAGGGGCTC	376
Qy	55	-----GluGlnAspArgProThrProSerProAlaProSerArgProPheSerVal	72
Db	375	ATGTGTCCAGCTGTCCAGACAGGTCCCAATCTGCTCCCATCTACCTGGGACATCCTG	316
Qy	73	---ArgAlaAsnAspValIleuThrLeuSerLeuThrAlaAlaGluTyrAspGlnSerThr	91
Db	315	TACAGGATTCCTCAGTCCTCTGG---GAACCTAACCCACCTGGTACCCATCTAGCTTC	259
Qy	92	TyrGlySerSer-----ThrGly	97
Db	258	TGGGGCTCCATCAGTCAGTGAGGAGTCACACCGGACCTGGACCCACCCCTGGG	199
Qy	98	ProValTyrValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaVal	117
Db	198	TCTGTCCATCTCAGTCCGGCTGAGCTCTGGGCAAGCCACCTCTTCTGAGCAGCA	139
Qy	118	AlaArgSerLeuAspTyrThrLysValThr	127
Db	138	GGCAGCGCAAGAGCTGGGAGCAGCAGACA	109

Search completed: October 28, 2004, 11:53:42  
Job time : 2044 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 28, 2004, 10:19:49 ; Search time 67 Seconds

(without alignments)  
2259.672 Million cell updates/sec

Title: US-10-089-292A-2

Perfect score: 1095

Sequence: 1 QLFYSRPVWSANGPTVKLY.....SLGAGPVSIASAVLAPPPR 213

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO spool/US10089292/runat 27102004 153928 10184/app query.fasta\_1.391  
-DB=Issued Patents NA -QFMT=fastap -SUPFIX=rni -MINMATCH=0.1 -LOOPCT=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pco -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10089292 @CGN 1 1 54 @runat 27102004 153928 10184 -NCPU=6 -ICPU=3  
-NO MMAP -LARGSEQURY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PTUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1076	98.3	803	4	US-09-172-699-17
2	1076	98.3	984	1	US-08-240-049B-7
3	1076	98.3	984	1	US-08-259-148A-7
4	1076	98.3	984	1	US-08-484-054-7
5	1076	98.3	984	2	US-07-876-941A-7
6	1076	98.3	984	3	US-08-542-634-5
7	1076	98.3	984	3	US-08-477-292-5
8	1076	98.3	984	4	US-07-870-985A-7
9	1076	98.3	984	5	PCT-US95-13703-5
10	1076	98.3	1311	1	US-08-259-148A-9
11	1076	98.3	1311	1	US-08-484-054-9
12	1076	98.3	1311	2	US-07-876-941A-9

13	1076	98.3	1311	4	US-07-870-985A-9	Sequence 9, Appli
14	1076	98.3	1647	3	US-08-542-634-3	Sequence 3, Appli
15	1076	98.3	1647	3	US-08-477-292-3	Sequence 3, Appli
16	1076	98.3	1647	5	PCT-US95-13703-3	Sequence 3, Appli
17	1076	98.3	1658	4	US-09-172-699-15	Sequence 15, Appli
18	1076	98.3	1686	4	US-09-172-699-19	Sequence 19, Appli
19	1076	98.3	1983	1	US-08-240-049B-1	Sequence 1, Appli
20	1076	98.3	1984	4	US-09-172-699-1	Sequence 1, Appli
21	1076	98.3	2049	1	US-08-259-148A-3	Sequence 3, Appli
22	1076	98.3	2049	1	US-08-484-054-3	Sequence 3, Appli
23	1076	98.3	2049	2	US-07-876-941A-3	Sequence 3, Appli
24	1076	98.3	2049	2	US-08-542-634-1	Sequence 1, Appli
25	1076	98.3	2049	3	US-08-477-292-1	Sequence 1, Appli
26	1076	98.3	2049	4	US-07-870-985A-3	Sequence 3, Appli
27	1076	98.3	2049	5	PCT-US95-13703-1	Sequence 1, Appli
28	1076	98.3	2094	1	US-08-259-148A-1	Sequence 1, Appli
29	1076	98.3	2094	1	US-08-484-054-1	Sequence 1, Appli
30	1076	98.3	2094	2	US-07-876-941A-1	Sequence 1, Appli
31	1076	98.3	2094	4	US-07-870-985A-1	Sequence 1, Appli
32	1076	98.3	7168	3	US-08-840-316-4	Sequence 4, Appli
33	1076	98.3	7168	3	US-08-809-523-4	Sequence 4, Appli
34	1076	98.3	7168	3	US-08-471-971-4	Sequence 4, Appli
35	1076	98.3	7168	4	US-09-402-776-4	Sequence 4, Appli
36	1076	98.3	7168	4	US-08-470-246-4	Sequence 4, Appli
37	1076	98.3	7168	4	US-08-316-765-4	Sequence 4, Appli
38	1076	98.3	7168	4	US-09-724-475-4	Sequence 4, Appli
39	1076	98.3	7168	5	PCT-US93-08849A-4	Sequence 4, Appli
40	1076	98.3	7168	5	PCT-US93-08849-4	Sequence 4, Appli
41	1076	98.3	7195	3	US-08-478-507-6	Sequence 6, Appli
42	1076	98.3	7195	3	US-09-128-275A-6	Sequence 6, Appli
43	1076	98.3	7195	3	US-09-553-427-6	Sequence 6, Appli
44	1043	95.3	981	1	US-08-259-148A-8	Sequence 8, Appli
45	1043	95.3	981	1	US-08-484-054-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1  
US-09-172-699-17  
; Sequence 17, Application US/09172699A  
; Patent No. 6514500  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David A.  
; APPLICANT: Locarnini, Stephen A.  
; APPLICANT: Tressi, Joseph  
; APPLICANT: Hui, Zhuang  
; APPLICANT: Li, Fan  
; TITLE OF INVENTION: IMMUNOREACTIVE ANTIGENS OF HEPATITIS E VIRUS  
; FILE REFERENCE: Davies Col. Cave  
; CURRENT APPLICATION NUMBER: US/09/172.699A  
; CURRENT FILING DATE: 1998-10-14  
; EARLIER APPLICATION NUMBER: 08/617,927  
; EARLIER FILING DATE: 1996-06-20  
; NUMBER OF SEQ ID NOS: 24  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 17  
; LENGTH: 803  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(801)  
US-09-172-699-17  
  
Alignment Scores:  
Pred. No.: 6.07e-122 Length: 803  
Score: 1076.00 Matches: 210  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.26% Indels: 0  
DB: 4 Gaps: 0  
  
US-10-089-292A-2 (1-213) x US-09-172-699-17 (1-803)

QY 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 20  
DB 1 CAGCTGTTCTACTCGTCCGCTGCTCAGCAATGGCAGCCGACTGTTAAGCTTTAT 60  
QY 21 ThrSerValGluAsnAlaGlnAspLysGlyIleAlaIleProHisAspLeu 40  
DB 61 ACATCTGTAGAAATGCTCAGCAGGATAAGGTAATTGCAATCCGCATGACATCGACCTC 120  
QY 41 GlyGluSerArgValIleGlnAspTyrAspAsnGlnHisGluAlaAspArgProThr 60  
DB 121 GGGAGTCTCGTGTAGTTATTGAGATTATGCAATGATGCTGTTGGCTT 180  
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrLeu 80  
DB 181 CTTTCCCGAGCCCATCGGCCCTTTCTGCTCCGAGCTAATGATGCTGTTGGCTT 240  
QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
DB 241 TCTCTCACCGCTGCGAGTATGACAGTCCACTTACGGCTCTTCGACGGCCCGCAGTCTAT 300  
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
DB 301 GTCCTGACTCTGTGACTCTGTGTTAATGTTGCGACCGCGCGAGCGGTTGCCCGGTCA 360  
QY 121 LeuAspTyrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140  
DB 361 CTCGAGTGGCAAGGTACACTTGATGTCGCCCTTCCACCATCCAGCAGTATCA 420  
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTyrPheGluAlaGlyThrThr 160  
DB 421 AAGACCTCTTGTCTCTGCGCTGCGGTAAGCTCTCTTTTGGGAGCAGGTACTACT 480  
QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180  
DB 481 AAAGCGGGTACCTTTAATTAATTAACACCTGCTAGTACCACTGCTGCTGTTGAGAA 540  
QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200  
DB 541 GCGCTGGGCATCGGTTGCTATTTCCACTTACACCACTAGCTAGCTGGTGTCTGCCGTC 600

## RESULT 2

US-08-240-049B-7

Sequence 7, Application US/08240049B

Patent No. 5686239

GENERAL INFORMATION:

APPLICANT: Reyes, Gregory R.

APPLICANT: Tam, Albert W.

APPLICANT: Yarbrough, Patrice O.

TITLE OF INVENTION: Hepatitis E Virus Peptides and Methods

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESS: Genelabs Technologies, Inc.

STREET: 505 Penobscot Drive

CITY: Redwood City

STATE: CA

COUNTRY: USA

ZIP: 94063

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/240,049B

FILING DATE: 09-MAY-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Charles K. Sholtz

REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 4600-0093.35; G32P4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 984 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain) SG3

US-08-240-049B-7

## Alignment Scores:

Pred. No.: 8.38e-122 Length: 984

Score: 1076.00 Matches: 210

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 98.26% Indels: 0

DB: 1 Gaps: 0

US-10-089-292A-2 (1-213) x US-08-240-049B-7 (1-984)

QY 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 20  
DB 181 CAGCTGTTCTACTCGGCTCCGCTGCTCAGCAATGGCAGCCGACTGTTAAGTGTAT 240  
QY 21 ThrSerValGluAsnAlaGlnAspLysGlyIleAlaIleProHisAspLeu 40  
DB 241 ACATCTGTAGAAATGCTCAGCAGGATAAGGTAATTGCAATCCGCGCATGACATTCACCTC 300  
QY 41 GlyGluSerArgValIleGlnAspTyrAspAsnGlnHisGluAlaAspArgProThr 60  
DB 301 GGAGAATCTCGTGTGTTTATTTCAGGATTAATGATTAACCAACATGAACAAGATCGGCCGAG 360  
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrLeu 80  
DB 361 CTTCTCCAGCCCCATCGGCCCTTCTCTGCTTCGAGCTAATGATGCTTTGGCTC 420  
QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
DB 421 TCTCTCACCGCTGCGAGTATGACAGTCCACTTATGCTCTTCGACTGGCCAGTTTAT 480  
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
DB 481 GTTCTGACTCTGTGACCTTGGTTAATGTTGCGACCGCGCGAGCGGCTTGCCCGGTG 540  
QY 121 LeuAspTyrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140  
DB 541 CTCGATTGGACCAAGGTACACTTTCAGCGTCCGCCCTCTCCACCATCCAGCAGTCTG 600  
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTyrPheGluAlaGlyThrThr 160  
DB 601 AAGACCTTCTTGTCTCGCGCTCCGCGTAAGCTCTCTTCTTGGGAGCAGCACAAC 660  
QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180  
DB 661 AAAGCGGGTACCTTTAATTAATTAACACCACTGTAGCAGCAACACTGCTGTCGAGAA 720  
QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200  
DB 721 GCGCGCGGACCGGGTCCGCTATTTCCACTTACACCACTAGCTGGTGGTGTCTGCCGTC 780  
QY 201 SerIleSerAlaValAlaValLeuAlaPro 210  
DB 781 TCCATTCTCGGGTTGCGGCTTTTAGCCCC 810

## RESULT 3

US-08-259-148A-7

Sequence 7, Application US/08259148A  
Patent No. 5741490  
GENERAL INFORMATION:  
APPLICANT: Reyes, Gregory R.  
APPLICANT: Bradley, Daniel W.  
APPLICANT: Twu, Jr-Shin  
APPLICANT: Purdy, Michael A.  
APPLICANT: Tam, Albert W.  
APPLICANT: Krawczynski, Krzysztof Z.  
APPLICANT: Yarbough, Patrice D.  
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/259,148A  
FILING DATE: 13-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 822,335  
FILING DATE: 17-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 505,888  
FILING DATE: 05-APR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 420,921  
FILING DATE: 13-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 367,486  
FILING DATE: 16-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 336,672  
FILING DATE: 11-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 208,997  
FILING DATE: 17-JUN-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 4600-0093.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 984 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: SG3, BURMA, FIGURE 7  
US-08-259-148A-7

Alignment Scores:  
Pred. No.: 8,38e-122 Length: 984  
Score: 1076.00 Matches: 210  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.26% Indels: 0  
Gaps: 1

US-10-089-292A-2 (1-213) x US-08-259-148A-7 (1-984)  
Qy 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 20  
Db 181 CAGCTGTTCTACTCCCGTCCCGTTGTTCTCAGCAATGGCGAGCGACTGTTAAGTTGAT 240  
Qy 21 ThrSerValGluAsnAlaGlnGlnAspLysGlyLeuAlaIleProHisAspLeu 40  
Db 241 ACATCTGTAGAGATGCTCAGCAGGATAAGGTTATTCGAATCCCGCATGACATTCACCTC 300  
Qy 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60  
Db 301 GGAGAACTCTGTGTTATTCAGGATTTATGATACCAACATGACACAGATCGCGCGAG 360  
Qy 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrLeu 80  
Db 361 CTTTCTCCAGCCCATCGCGCTTTCTCTGCTTCGAGCTAATGATGTGCTTTGGCTC 420  
Qy 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
Db 421 TCTCTCACCCTGCCGAGTATGACCACTTATGCTCTTCGACTGGCCCACTTTAT 480  
Qy 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyValAlaGlnAlaValAlaArgSer 120  
Db 481 GTTCTGACTCTGTGACCTTGTTAAATGTTGGACCGCGCGAGCGCTTCCCGCTCG 540  
Qy 121 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrThrIleGlnGlnTyrSer 140  
Db 541 CTCGATTGGACCAAGGTCACACTTGACGTCGCCCTCTCCACCATCCAGCAGTACTCG 600  
Qy 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160  
Db 601 AAGACCTTCTTTGTCCTGCCCTCCGCGTAAAGTCTCTTTCTGGGAGCGGCAACT 660  
Qy 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180  
Db 661 AAAGCGGTACCTTATATATATACCACTGCTAGCAGCAACTCTCTTGTCCAGAT 720  
Qy 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200  
Db 721 GCCCGCGGACCGCGTTCGCTATTTCCACTTACCACTAGCTAGCTGGTGGTGGTCCCGTC 780  
Qy 201 SerLeuSerAlaValAlaValLeuAlaPro 210  
Db 781 TCCATTTCTGCGGTTCGCTTATGACCCCT 810  
RESULT 4  
US-08-484-054-7  
Sequence 7, Application US/08484054  
Patent No. 5770689  
GENERAL INFORMATION:  
APPLICANT: Reyes, Gregory R.  
APPLICANT: Bradley, Daniel W.  
APPLICANT: Twu, Jr-Shin  
APPLICANT: Purdy, Michael A.  
APPLICANT: Tam, Albert W.  
APPLICANT: Krawczynski, Krzysztof Z.  
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,054

aa against 0.00

1

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 984 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: SG3, BURMA, FIGURE 7  
US-07-876-941A-7

Alignment Scores:  
Pred. No.: 8,38e-122 Length: 984  
Score: 1076.00 Matches: 210  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.26% Indels: 0  
DB: 2 Gaps: 0

US-10-089-292A-2 (1-213) x US-07-876-941A-7 (1-984)

```
Qy 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValIysLeuTyr 20
Db 181 CAGCTGTTCTACTCCCGTCCCGTCTCTCAGCAATGGCGAGCGACTGTTAAGTTGTAT 240
Qy 21 ThrSerValGluAsnAlaGlnAspIysGlyLeuAlaIleProHisAspIleAspLeu 40
Db 241 ACATCTGTAGAGATGCTCAGCAGGATAGAGGTATTGCAATCCCGCATGACATTGACCTC 300
Qy 41 GlyGluSerArgValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60
Db 301 GGAGAAATCTCGTGTGTTATTTCAGGATTTATGTAACCAACATGTAACCAAGATCGGCCGACG 360
Qy 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyr 80
Db 361 CCTTCTCCAGCCCATCGCGCCCTTCTCTGCTCTCGAGCTAATGATGCTTTGGCTC 420
Qy 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100
Db 421 TCTCTCACCGTCCGAGTATGACAGTCCATTTATGCTCTTCGACTGGCCCACTTTAT 480
Qy 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120
Db 481 GTTCTGACTCTGTGACCTTGGTTAATGTTGGACCGCGCGCAGCGCGTGGCCGGTTCG 540
Qy 121 LeuAspThrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140
Db 541 CTCGATTGGACCAAGTCCACACTTGACGGTCCGCCCTCTCCACCATCCAGCAGTACTCG 600
Qy 141 LysThrPhePheValLeuProLeuArgGlyIysLeuSerPheTyrGluAlaGlyThrThr 160
Db 601 AAGACCTTCTTTGTCCTCCCGCTCCCGGTAGCTCTCTTTTGGAGAGGAGGACAACT 660
Qy 161 LysAlaGlyTyrProTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180
Db 661 AAAGCCGGTACCTTATATATATATACCACTGCTAGCACCACACTCTCTTGTTCGAGAA 720
Qy 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200
Db 721 GCCCGCGGCGCCGGTCCGCTATTTCACCTTACACCACTAGCCCTGGGTGCTGCCCGTC 780
Qy 201 SerIleSerAlaValAlaValLeuAlaPro 210
Db 781 TCCATTTCTGCGGTTCGCGTTTATAGCCCCC 810
```

RESULT 6

US-08-542-634-5

; Sequence 5, Application US/08542634

; Patent No. 6214970

GENERAL INFORMATION:  
APPLICANT: Fuerst, Thomas R.  
APPLICANT: Mcatee, C. Patrick  
APPLICANT: Yarbough, Patrice O.  
APPLICANT: Zhang, Yifan  
TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Ave., Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/542,634  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 4600-0293.30  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 984 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain) SG3  
INDIVIDUAL ISOLATE: region  
US-08-542-634-5

Alignment Scores:  
Pred. No.: 8,38e-122 Length: 984  
Score: 1076.00 Matches: 210  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.26% Indels: 0  
DB: 3 Gaps: 0

US-10-089-292A-2 (1-213) x US-08-542-634-5 (1-984)

```
Qy 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValIysLeuTyr 20
Db 181 CAGCTGTTCTACTCCCGTCCCGTCTCTCAGCAATGGCGAGCGACTGTTAAGTTGTAT 240
Qy 21 ThrSerValGluAsnAlaGlnAspIysGlyLeuAlaIleProHisAspIleAspLeu 40
Db 241 ACATCTGTAGAGATGCTCAGCAGGATAGAGGTATTGCAATCCCGCATGACATTGACCTC 300
Qy 41 GlyGluSerArgValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60
Db 301 GGAGAAATCTCGTGTGTTATTTCAGGATTTATGTAACCAACATGTAACCAAGATCGGCCGACG 360
Qy 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyr 80
Db 361 CCTTCTCCAGCCCATCGCGCCCTTCTCTGCTCTCGAGCTAATGATGCTTTGGCTC 420
Qy 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100
Db 421 TCTCTCACCGTCCGAGTATGACAGTCCACTTATGCTTATGCTCTTCGACTGGCCCACTTTAT 480
```





COUNTRY: USA  
 ZIP: 94306  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 FILING DATE: 20-APRIL-1992  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 822,335  
 FILING DATE: 17-JAN-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 505,888  
 FILING DATE: 05-APRIL-1990  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 420,921  
 FILING DATE: 13-OCTOBER-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 367,486  
 FILING DATE: 16-JUNE-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 336,672  
 FILING DATE: 11-APRIL-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 208,997  
 FILING DATE: 17-JUNE-1988  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Fabian, Gary R.  
 REGISTRATION NUMBER: 33,875  
 REFERENCE/DOCKET NUMBER: 4500-0093.30  
 TELEPHONE: (415) 324-0880  
 TELEFAX: (415) 324-0960  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 984 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA to mRNA  
 HYPOTHETICAL: NO  
 ORIGINAL SOURCE:  
 INDIVIDUAL ISOLATE: SG3, BURMA, FIGURE 7  
 US-07-870-985A-7

Alignment Scores:  
 Pred. No.: 8.38e-122 Length: 984  
 Score: 1076.00 Matches: 210  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.26% Indels: 0  
 DB: 4 Gaps: 0

US-10-089-292A-2 (1-213) x US-07-870-985A-7 (1-984)

Qy 1 GlnLeuPheTyrSerArgProValValSerAlaHengGlyGluProThrValLysLeuTyr 20  
 Db 181 CAGCTGTTCTACTCCCGCCGTTGCTCAGCCAAATGGCGAGCCGACTGTTAAGTTGAT 240  
 Qy 21 ThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 40  
 Db 241 ACATCTGTAGAGATGCTCAGCAGGATAGGGTATTGCAATCCCGCATGACATTGACCTC 300  
 Qy 41 GlyGluSerArgValValIleGlnAspTyrAspGlnHisGluGlnAspArgProThr 60  
 Db 301 GGAGAATCTCGTGGTTATTTCAGGATTATGATACCAACATGAACAAGATCGGCCGACG 360  
 Qy 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrLeu 80  
 Db 361 CCTTCTCCAGCCCATCGCCGCTTCTCTGCTTCGAGCTAATGATGCTGCTTGGCTC 420

Qy 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
 Db 421 TCTCTCACCGCTGCGAGTATGACAGTCCACTTATGGCTCTTCGACTGGCCAGTTTAT 480  
 Qy 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
 Db 481 GTTTCGTACTCTGTGACCTTGGTTAATGTTGCGACCGCGCGAGCGCTTGGCCGGTCG 540  
 Qy 121 LeuAspTyrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140  
 Db 541 CTCGATTGGACCAAGGTCACACTTTCGCGGTGCGCCCTCTCCACCATCCAGCAGTACTCG 600  
 Qy 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160  
 Db 601 AAGACCTTCTTTGTCTGCTGCGCTCCGCGGTAAAGCTCTCTTCTGGGAGGCGACCAACT 660  
 Qy 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluLeu 180  
 Db 661 AAGCCCGGTACCTTATATTAATTAACACCACTGCTAGCAGCAACTGCTTGTGCGAAT 720  
 Qy 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProVal 200  
 Db 721 GCCCGCGGACCGCGTGGCTATTTCCACTTACACCACTAGCTGGGTGGTGGTCCGCTC 780  
 Qy 201 SerIleSerAlaValAlaValLeuAlaPro 210  
 Db 781 TCCATTCTCGCGTTGCGGTTTAGCCCCC 810

RESULT 9  
 PCT-US95-13703-5  
 ; Sequence 5, Application PC/TUS9513703  
 ; GENERAL INFORMATION:  
 ; APPLICANT:  
 ; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND  
 ; TITLE OF INVENTION: USES THEREFOR  
 ; NUMBER OF SEQUENCES: 31  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Dehlinger & Associates  
 ; STREET: P.O. Box 60850  
 ; CITY: Palo Alto  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94306-0850  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patent in Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US95/13703  
 ; FILING DATE:  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Fabian, Gary R.  
 ; REGISTRATION NUMBER: 33,875  
 ; REFERENCE/DOCKET NUMBER: 4500-0293.41  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 324-0880  
 ; TELEFAX: (415) 324-0960  
 ; INFORMATION FOR SEQ ID NO: 5:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 984 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; HYPOTHETICAL: NO  
 ; ORIGINAL SOURCE:  
 ; INDIVIDUAL ISOLATE: Hepatitis E Virus (Burma strain) SG3  
 ; INDIVIDUAL ISOLATE: region  
 ; PCT-US95-13703-5

Mon Nov 1 10:49:17 2004

us-10-089-292a-2.rni

Alignment Scores: 8,38e-122 Length: 984  
Pred. No.: 1076.00 Matches: 210  
Score: 100.00% Conservatives: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 98.26% Gaps: 0  
DB: 5

US-10-089-292A-2 (1-213) x PCT-US95-13703-5 (1-984)

QY 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValValLeuTyr 20  
DB 181 CAGCTGTTCTACTCCGCTCCGCTTCTCAGCAATGGCGAGCGACTGTTAAGTTGAT 240  
QY 21 ThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 40  
DB 241 ACATCTGTAGAAATGCTCAGCAGATAGAGGTATTCATCCCGCATGACATTGACCTC 300  
QY 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60  
DB 301 GGAGAAATCTCGTGTGTTATTCAGGATATGATAACCAACATGAACAAGATCGCGCAGC 360  
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrLeu 80  
DB 361 CCTTCTCCAGCCCATCGCGCTTCTCTGCTTCGAGCTAATGATGCTTGGCTC 420  
QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
DB 421 TCTCTCAGCGTCCGAGTATGACCACTTATGCTCTTCGACTGGCCAGTTTAT 480  
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
DB 481 GTTCTGACTCTGTGACCTTGGTAAATGTGGACCGCGCGAGCGCTTCCCGTGC 540  
QY 121 LeuAspThrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140  
DB 541 CTGATTTGGACCAAGTCACTACCTTACGCTGCGCTTCCCTTCCACCATCCAGCAGTACG 600  
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThr 160  
DB 601 AAGACCTTCTTGTCTGCGCTCCGCGTAAAGCTCTCTTCTGGAGGAGGAGCAACT 660  
QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsn 180  
DB 661 AAAGCCGGTACCTTATATATATATACACCACTGCTAGCAGCAACTGCTTGTGAGAAT 720  
QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200  
DB 721 GCCGCGCGGACCGGCTCGCTATTTCCACTTACACCACTAGCCTGGTGTGCTGCCGTC 780  
QY 201 SerLeuSerAlaValAlaValLeuAlaPro 210  
DB 781 TCCATTTCTGCGGTTCGCGTTTATAGCCCCC 810

RESULT 10

US-08-259-148A-9  
Sequence 9, Application US/08259148A  
Patent No. 5741490

GENERAL INFORMATION:

APPLICANT: Reyes, Gregory R.  
APPLICANT: Bradley, Daniel W.  
APPLICANT: Twu, Jr-Shin  
APPLICANT: Purdy, Michael A.  
APPLICANT: Tam, Albert W.  
APPLICANT: Krawczynski, Krzysztof Z.  
APPLICANT: Yarborough, Patricia D.  
TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method  
NUMBER OF SEQUENCES: 60  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA

COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/259,148A  
FILING DATE: 13-JUN-1994  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 822,335  
FILING DATE: 17-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 505,888  
FILING DATE: 05-APR-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 420,921  
FILING DATE: 13-OCT-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 367,486  
FILING DATE: 16-JUN-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 336,672  
FILING DATE: 11-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 208,997  
FILING DATE: 17-JUN-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Sholtz, Charles K.  
REGISTRATION NUMBER: 38,615  
REFERENCE/DOCKET NUMBER: 4600-0093.20  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1311 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: C2, BURMA, FIGURE 7

US-08-259-148A-9

Alignment Scores:  
Pred. No.: 1,32e-121 Length: 1311  
Score: 1076.00 Matches: 210  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.26% Indels: 0  
DB: 1 Gaps: 0

US-10-089-292A-2 (1-213) x US-08-259-148A-9 (1-1311)

QY 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValValLeuTyr 20  
DB 508 CAGCTGTTCTACTCCGCTCCGCTTCTCAGCAATGGCGAGCGACTGTTAAGTTGAT 567  
QY 21 ThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 40  
DB 568 ACATCTGTAGAAATGCTCAGCAGATAGAGGTATTCATCCCGCATGACATTGACCTC 627  
QY 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60  
DB 628 GGAGAAATCTCGTGTGTTATTCAGGATATGATAACCAACATGAACAAGATCGCGCAGC 687  
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrLeu 80  
DB 688 CCTTCTCCAGCCCATCGCGCTTCTCTGCTTCGAGCTAATGATGCTTGTGCTC 747

QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
DB 748 TCTCTCAGCGTGGCGAGTATGACCACTTATGCTTCTTCAGTGGCCAGTTTAT 807  
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
DB 808 GTTTCCTGACTCTGTGACCTTGTGTTAATGTTGCGACCGCGCGAGCGGTTGCCCGGTG 867  
QY 121 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140  
DB 868 CTCGATTGACCAAGTCCACTTGACGGTGGCCCTCTCCACCATCCAGCAGTACTCG 927  
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160  
DB 928 AAGACCTTCTTTGTCTCGCGTCCGCGTAAAGTCTCTTTCTGGGAGGCGACCAACT 987  
QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180  
DB 988 AAAGCGGGTACCCCTTAAATTAATTAACACCACTGCTAGCACAACCTGTTGTCGAGAA 1047  
QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200  
DB 1048 GCCGCGGCGACCGCGTGCCTATTTCCACTTACACCACTAGCTGGTGGTGGTCCGCTC 1107  
QY 201 SerIleSerAlaValAlaValLeuAlaPro 210  
DB 1108 TCCATTTCGCGGTGGCGGTTTGTAGCCGCC 1137  
RESULT 11  
US-08-484-054-9  
; Sequence 9, Application US/08484054  
; Patent No. 5770689  
; GENERAL INFORMATION:  
; APPLICANT: Reyes, Gregory R.  
; APPLICANT: Bradley, Daniel W.  
; APPLICANT: Twu, Jr-Shin  
; APPLICANT: Purdy, Michael A.  
; APPLICANT: Tam, Albert W.  
; APPLICANT: Krawczynski, Krzysztof Z.  
; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Avenue, Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,054  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 870,985  
; FILING DATE: 20-APRIL-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 822,335  
; FILING DATE: 17-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 505,888  
; FILING DATE: 05-APRIL-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 420,921  
; FILING DATE: 13-OCTOBER-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 367,486  
; FILING DATE: 16-JUNE-1989

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 336,672  
; FILING DATE: 11-APRIL-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 208,997  
; FILING DATE: 17-JUNE-1988  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sholtz, Charles K.  
; REGISTRATION NUMBER: 38,615  
; REFERENCE/DOCKET NUMBER: 4600-0093.38  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1311 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA to mRNA  
; HYPOTHEetical: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: C2, BURMA, FIGURE 7  
US-08-484-054-9

Alignment Scores:  
Pred. No.: 1,32e-121 Length: 1311  
Score: 1076.00 Matches: 210  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.26% Indels: 0  
DB: 1 Gaps: 0

US-10-089-292A-2 (1-213) x US-08-484-054-9 (1-1311)

QY 1 GlnLeuPheTyrSerArgProValValSerAlaGlnGlyProThrValLysLeuTyr 20  
DB 508 CAGCTGTGTACTCCGCTCCGTTGTCTCAGCCCAATGGCAGCGGACTGTTAAGTTGTAT 567  
QY 21 ThrSerValGluAsnAlaGlnAlaGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 40  
DB 568 ACATCTGTAGAGATGCTCAGCAGGATTAAGGTATTGCAATCCCGCATGACATTGACCTC 627  
QY 41 GlyLysSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60  
DB 628 GGAGAACTCTCGTGTGTTATTTCAGGATTATGATAACCAACATGAACAGATCGGCGGACG 687  
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80  
DB 688 CTTTCTCAGCCCATCGGCCCTTTCTCTGCTCTCGAGCTAATGATGTGCTTTGGCTC 747  
QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
DB 748 TCTCTCAGCGTGGCGAGTATGACCACTTATGCTTATGCTTCTTCAGTGGCCAGTTTAT 807  
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
DB 808 GTTTCCTGACTCTGTGACCTTGTGTTAATGTTGCGACCGCGCGAGCGGTTGCCCGGTG 867  
QY 121 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140  
DB 868 CTCGATTGACCAAGTCCACTTGACGGTGGCCCTCTCCACCATCCAGCAGTACTCG 927  
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160  
DB 928 AAGACCTTCTTTGTCTCGCGTAAAGTCTCTTTCTGGGAGGCGACCAACT 987  
QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180  
DB 988 AAAGCGGGTACCCCTTAAATTAATTAACACCACTGCTAGCACAACCTGTTGTCGAGAA 1047  
QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200

Db 1048 GCCGCGGCGCAGCGGTGCGCTATTTCCACTTACACCACTAGCTGGGTGCTGGTCCCGTC 1107  
Qy 201 SerlleSerAlaValAlaValAlaPro 210  
Db 1108 TCCATTTCTGCGGTGCGGTTTGTAGCCGCC 1137

## RESULT 12

US-07-876-941A-9

; Sequence 9, Application US/07876941A

; Patent No. 5885768

; GENERAL INFORMATION:

; APPLICANT: Reyes, Gregory R.

; APPLICANT: Bradley, Daniel W.

; APPLICANT: Tam, Albert W.

; APPLICANT: Mitchell, Carl

; TITLE OF INVENTION: Hepatitis E Virus Peptide Antigen and

; TITLE OF INVENTION: Antibodies

; NUMBER OF SEQUENCES: 76

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dehlinger &amp; Associates

; STREET: 350 Cambridge Avenue, Suite 250

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/876,941A

; FILING DATE: 01-MAY-1992

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 822,335

; FILING DATE: 17-JAN-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 505,888

; FILING DATE: 05-APRIL-1990

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 420,921

; FILING DATE: 13-OCTOBER-1989

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 367,486

; FILING DATE: 16-JUNE-1989

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 336,672

; FILING DATE: 11-APRIL-1989

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 208,997

; FILING DATE: 17-JUNE-1988

; ATTORNEY/AGENT INFORMATION:

; NAME: Sholtz, Charles K.

; REGISTRATION NUMBER: 38,615

; REFERENCE/DOCKET NUMBER: 4600-0093.33

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 324-0880

; TELEFAX: (415) 324-0960

; INFORMATION FOR SEQ ID NO: 9:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1311 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA to mRNA

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; INDIVIDUAL ISOLATE: C2, BURMA, FIGURE 7

; US-07-876-941A-9

; Alignment Scores:

; Pred. No.: 1.32e-121 Length: 1311

Score: 1076.00 Matches: 210  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.26% Indels: 0  
DB: 2 Gaps: 0

US-10-089-292A-2 (1-213) x US-07-876-941A-9 (1-1311)

Qy 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 20  
Db 508 CAGCTGTTCTACTCCCGTCCGTTGCTCAGCAATGCGGAGCGGACTGTTAAGTTGAT 567  
Qy 21 ThrSerValGluAsnAlaGlnAspGlyGlyLeuAlaLeuProHisAspLeu 40  
Db 568 ACATCTGTAGAAATGCTCAGCAGATAAGGGTATTGCAATCCCGCATGATTGAGCTC 627  
Qy 41 GlyGluSerArgValValLeuGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60  
Db 628 GGAGAATCTCGTGGGTATTTCAGGATTATGATACCAACATGAAACAGATCGGCCGACG 687  
Qy 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrLeu 80  
Db 688 CTTCTCCAGCCCATCGCGCCCTTTCTCTGCTCTCGAGCTAATGATGCTTTGGCTC 747  
Qy 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
Db 748 TCTCTCACCGCTGCGGAGTATGACAGCTCACTTATGGCTCTTCAGCTGCGCCGATTTAT 807  
Qy 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
Db 808 GTTTCAGCTCTGTGACCTTGTTGTTAATGTTGCGACCGGCGGCGGAGCGGTTGCCCGTGC 867  
Qy 121 LeuAspTyrThrLysValThrLeuAspGlyArgProLeuSerThrLeuGlnTyrSer 140  
Db 868 CTCGATTGACCAAGGTCACACTTGACGGTCGCCCTCTCCACCATCCAGCAGTACTCG 927  
Qy 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThr 160  
Db 928 AAGACCTTTCTTTGCTCGCGCTCGCGGTAAAGTCTCTTTCTGGGAGGAGGACCAACT 987  
Qy 161 LysAlaGlyTyrProTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180  
Db 988 AAAGCCGGGTACCTTATTAATTAACACCACTGCTAGCGACCAACTGCTTTGTGAGAAT 1047  
Qy 181 AlaAlaGlyHisArgValAlaLeuSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200  
Db 1048 GCCGCGGCGCAGCGGTGCTGCTATTTCCACTTACACCACTAGCTGGGTGCTGGTCCCGTC 1107  
Qy 201 SerlleSerAlaValAlaValAlaPro 210  
Db 1108 TCCATTTCTGCGGTGCGGTTTGTAGCCGCC 1137

## RESULT 13

US-07-870-985A-9

; Sequence 9, Application US/07870985A

; Patent No. 6455492

; GENERAL INFORMATION:

; APPLICANT: Reyes, Gregory R.

; APPLICANT: Bradley, Daniel W.

; APPLICANT: Twu, Jr-Shin

; APPLICANT: Purdy, Michael A.

; APPLICANT: Tam, Albert W.

; APPLICANT: Krawczynski, Krzysztof Z.

; TITLE OF INVENTION: Hepatitis E Virus Vaccine and Method

; NUMBER OF SEQUENCES: 22

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dehlinger &amp; Associates

; STREET: 350 Cambridge Avenue, Suite 250

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICANT: Fabian, Gary R.  
FILING DATE: 20-APRIL-1992  
FILING DATE: 20-APRIL-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 822,335  
FILING DATE: 17-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 505,888  
FILING DATE: 05-APRIL-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 420,921  
FILING DATE: 13-OCTOBER-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 367,486  
FILING DATE: 16-JUNE-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 336,672  
FILING DATE: 11-APRIL-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 208,997  
FILING DATE: 17-JUNE-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: Fabian, Gary R.  
REGISTRATION NUMBER: 33,875  
REFERENCE/DOCKET NUMBER: 4600-0093.30  
TELEPHONE: (415) 324-0880  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1311 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA to mRNA  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: C2, BURMA, FIGURE 7

US-07-870-985A-9

Alignment Scores:

Pred. No.: 1,328-121 Length: 1311

Score: 1076.00 Matches: 210

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 98.26% Indels: 0

DB: 4 Gaps: 0

US-10-089-292A-2 (1-213) x US-07-870-985A-9 (1-1311)

Qy 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 20

Db 508 CAGCTGTTCTACTCCGTCCTGCTCTCAGCAATGGCGAGCCGACTGTAAGTTGAT 567

Qy 21 ThrSerValGluAsnAlaGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 40

Db 568 ACATCTGTAGAGATGCTCAGCAGGATAAGGTAATGCAATCCCGCATGACATTGACCTC 627

Qy 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60

Db 628 GGAGAATCTCGTGTGTTATTTCAGATTATGATAACCAACATGAACAGATCGGCCGACG 687

Qy 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrLeu 80

Db 688 CTTCTCCAGCCCATCGGCCCTTCTCTGCTCTTCAGCTAATGATGCTTTGGCTC 747

Qy 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100

Db 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100

Db 748 TCTCTCACCCTGCCGAGTATGACCAGTCACCTTATGGCTCTTCGACTGGCCAGTTTAT 807

Qy 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120

Db 808 GTTTCCTGACTCTGTGACCTTGGTTAATGTTGCGAGCCGCGCAGGCGGTTGCCGCTCG 867

Qy 121 LeuAspTyrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140

Db 868 CTCGATTGGACCAAGGTACACTTGGAGGTGCGCCCTCTCCACCATCCACGACTACTCG 927

Qy 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThr 160

Db 928 AAGACCTTCTTTGCTCCGCTCCGCGGTAAAGCTCTCTTCTGGAGGAGGACCAACT 987

Qy 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180

Db 988 AAGCCGGGTACCTTATAATTAACACCACTGCTAGCGACCAACTGCTTGTGCGAGAT 1047

Qy 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200

Db 1048 GCCCGCGGACCGGTCGCTATTTCACCTTACACCACTAGCTGGGTGCTGGTCCCGTC 1107

Qy 201 SerIleSerAlaValAlaValLeuAlaPro 210

Db 1108 TCCATTCTCGCGTTGCCGTTTACGCCCC 1137

RESULT 14

US-08-542-634-3

; Sequence 3, Application US/08542634

; Patent No. 6214970

; GENERAL INFORMATION:

; APPLICANT: Fuerst, Thomas R.

; APPLICANT: McAtee, C. Patrick

; APPLICANT: Yarbough, Patricia O.

; APPLICANT: Zhang, Yifan

; TITLE OF INVENTION: HEPATITIS E VIRUS ANTIGENS AND USES THEREFOR

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dehlinger & Associates

; STREET: 350 Cambridge Ave., Suite 250

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/542,634

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Fabian, Gary R.

; REGISTRATION NUMBER: 33,875

; REFERENCE/DOCKET NUMBER: 4600-0293.30

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 324-0880

; TELEFAX: (415) 324-0960

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1647 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; INDIVIDUAL ISOLATE: Hepatitis E virus (Burma) r62kDa,

; INDIVIDUAL ISOLATE: FIGURE 2

; US-08-542-634-3

## Alignment Scores:

Pred. No.: 1,9e-121 Length: 1647  
Score: 1076.00 Matches: 210  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.26% Indels: 0  
DB: 3 Gaps: 0

US-10-089-292A-2 (1-213) x US-08-542-634-3 (1-1647)

QY 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 20  
DB 847 CAGCTGTTCTACTCCGCTCCGCTTCTCAGCAATGCGCAGCGACTGTTAAGTTGTAT 906  
QY 21 ThrSerValGluAsnAlaGlnAspLysGlyIleAlaProHisAspLeu 40  
DB 907 ACATCTGTAGAAATGCTCAGCAGGATAAGGTAATGCAATCCGATGACATTCACCTC 966  
QY 41 GlyUserArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60  
DB 967 GCAGAAATCTCGTGTGTTATTGAGATTATGATACCAACATGAACAGATCGGCCGAG 1026  
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80  
DB 1027 CCTTCTCCAGCCCATCGCGCCCTTCTCTGCTCTCGAGCTAATGCTTTGGCTC 1086  
QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerThrGlyProValTyr 100  
DB 1087 TCTCTCACCGCTGCGAGTATGACCGTCCATCTTATGGCTCTTCGACTGGCCAGTTAT 1146  
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
DB 1147 GTTCTGACTCTGTGACCTTGTGTTATGTCAGCGCGCGCAGCGCTTCCCGGTG 1206  
QY 121 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140  
DB 1207 CTCGATTGACCAAGTCACTTACGCGTCCGCCCTCTCCACCATCCAGCAGTACTCG 1266  
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160  
DB 1267 AAGACCTTCTTGTCTCGCGTTCGCGGTAGCTCTCTTTCTGGAGCAGGCAACT 1326  
QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180  
DB 1327 AAAGCGGGTACCTTATATTAATTAACACCACTGCTAGGACCACTGCTTCGAGAA 1386  
QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProVal 200  
DB 1387 GCCGCGGGCACCGGTCGCTATTTCCACTTACCCACTAGCCTAGCGTGGTCTCCCGTC 1446  
QY 201 SerIleSerAlaValAlaValLeuAlaPro 210  
DB 1447 TCCATTTCTGCGGTGCGGTTTTAGCCCCC 1476

## RESULT 15

US-08-477-292-3

Sequence 3, Application US/08477292

Patent No. 6291641

## GENERAL INFORMATION:

APPLICANT: Fuerst, Thomas

APPLICANT: McAtee, Patrick

APPLICANT: Yarbough, Patrice

APPLICANT: Zhang, Yifan

TITLE OF INVENTION: Hepatitis E Virus Antigens and Methods

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cathleen M. Desjardins, M.D.

STREET: 505 Penobscot Drive

CITY: Redwood City

STATE: CA

COUNTRY: US

ZIP: 94063

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,292  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/327,952  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brookes, Allan A  
REGISTRATION NUMBER: 36,373  
REFERENCE/DOCKET NUMBER: G32P5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 368-9500  
TELEFAX: (415) 368-0709  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1647 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: Hepatitis E virus (Burma) r62kDa,  
INDIVIDUAL ISOLATE: FIGURE 2

US-08-477-292-3

## Alignment Scores:

Pred. No.: 1,9e-121 Length: 1647  
Score: 1076.00 Matches: 210  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.26% Indels: 0  
DB: 3 Gaps: 0

US-10-089-292A-2 (1-213) x US-08-477-292-3 (1-1647)

QY 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 20  
DB 847 CAGCTGTTCTACTCCGCTCCGCTTCTCAGCAATGCGCAGCGACTGTTAAGTTGTAT 906  
QY 21 ThrSerValGluAsnAlaGlnAspLysGlyIleAlaProHisAspLeu 40  
DB 907 ACATCTGTAGAAATGCTCAGCAGGATAAGGTAATGCAATCCGCGCATGACATTCACCTC 966  
QY 41 GlyUserArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60  
DB 967 GCAGAAATCTCGTGTGTTATTGAGATTATGATACCAACATGAACAGATCGGCCGAG 1026  
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80  
DB 1027 CCTTCTCCAGCCCATCGCGCCCTTCTCTGCTCTCGAGCTAATGCTTTGGCTC 1086  
QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerThrGlyProValTyr 100  
DB 1087 TCTCTCACCGCTGCGAGTATGACCGTCCATCTTATGGCTCTTCGACTGGCCAGTTAT 1146  
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
DB 1147 GTTCTGACTCTGTGACCTTGTGTTATGTCAGCGCGCGCAGCGCTTCCCGGTG 1206  
QY 121 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140  
DB 1207 CTCGATTGACCAAGTCACTTACGCGTCCGCCCTCTCCACCATCCAGCAGTACTCG 1266  
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160  
DB 1267 AAGACCTTCTTGTCTCGCGTTCGCGGTAGCTCTCTTTCTGGAGCAGGCAACT 1326

Qy 161 LysAlaGlyTyrProTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||  
1327 AAAGCCGGGTACCCCTATAATTATAACACCACTGCTAGCGACCAACTGCTTGTGAGAAT 1386  
Qy 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200  
Db ||||||||||||||||||||||||||||||||||||||||||||||||  
1387 GCCGCCGGGCACCGGGTCGCTATTTCACCTTACACCACTAGCCTGGGTGCTGGTCCCGTC 1446  
Qy 201 SerIleSerAlaValAlaValLeuAlaPro 210  
Db ||||||||||||||||||||||||||||||||||||||||||||  
1447 TCCATTTCTGCGGTTGCCGTTTTAGCCCCC 1476

Search completed: October 28, 2004, 11:55:03  
Job time : 73 secs

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 28, 2004, 09:51:24 ; Search time 2714 Seconds  
(without alignments)  
3711.388 Million cell updates/sec

Title: US-10-089-292A-2  
Perfect score: 1095  
Sequence: 1 QLFYSRPVSVANGSPVTKLY.....SLGAGPVSVISAVVAPP 213

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO\_epool/US10089292/runat\_27102004\_153927\_10154/app\_query.fasta\_1.391  
-DB=GenEmbl -QWMT=fastap -SUPFIX=rge -MINMATCH=0.1 -LOECL=0 -LOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US10089292 @CGN 1 1 2527 @runat\_27102004\_153927\_10154 -NCPU=6 -ICPU=3  
-NO.MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*

- 1: gb.ba.\*
- 2: gb.htg.\*
- 3: gb.in.\*
- 4: gb.om.\*
- 5: gb.ov.\*
- 6: gb.pat.\*
- 7: gb.ph.\*
- 8: gb.pl.\*
- 9: gb.pr.\*
- 10: gb.ro.\*
- 11: gb.sts.\*
- 12: gb.sv.\*
- 13: gb.un.\*
- 14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1076	98.3	803	6	AR278919 Sequence
2	1076	98.3	984	6	AR002164 Sequence
3	1076	98.3	984	6	AR170408 Sequence
4	1076	98.3	984	6	I73145 Sequence 7

5	1076	98.3	984	6	AR232469	AR232469 Sequence
6	1076	98.3	1117	14	HEU40044	U40044 Hepatitis E
7	1076	98.3	1311	6	AR002166	AR002166 Sequence
8	1076	98.3	1311	6	AR232471	AR232471 Sequence
9	1076	98.3	1647	6	AR170406	AR170406 Sequence
10	1076	98.3	1647	6	E17105	E17105 DNA encodin
11	1076	98.3	1658	6	AR278918	AR278918 Sequence
12	1076	98.3	1686	6	AR278920	AR278920 Sequence
13	1076	98.3	1946	14	HEVBCN6	AF058684 Hepatitis E
14	1076	98.3	1983	6	E17107	E17107 Hepatitis E
15	1076	98.3	1983	6	I73139	I73139 Sequence 1
16	1076	98.3	1984	6	AR278906	AR278906 Sequence
17	1076	98.3	2049	6	AR002160	AR002160 Sequence
18	1076	98.3	2049	6	AR170404	AR170404 Sequence
19	1076	98.3	2049	6	AR232465	AR232465 Sequence
20	1076	98.3	2094	6	AR002158	AR002158 Sequence
21	1076	98.3	2094	6	AR232463	AR232463 Sequence
22	1076	98.3	2173	14	AF065061	AF065061 Hepatitis E
23	1076	98.3	2173	14	HPEA	M80581 Hepatitis E
24	1076	98.3	7138	14	AF185822	AF185822 Hepatitis E
25	1076	98.3	7168	6	AR139826	AR139826 Sequence
26	1076	98.3	7168	6	AR167470	AR167470 Sequence
27	1076	98.3	7168	6	AR234194	AR234194 Sequence
28	1076	98.3	7168	6	AR476111	AR476111 Sequence
29	1076	98.3	7168	6	AR487996	AR487996 Sequence
30	1076	98.3	7168	6	BD084498	BD084498 Recombina
31	1076	98.3	7194	6	E17109	E17109 Hepatitis E
32	1076	98.3	7195	6	AR150915	AR150915 Sequence
33	1076	98.3	7195	6	AR207633	AR207633 Sequence
34	1076	98.3	7204	14	AF444002	AF444002 Hepatitis E
35	1076	98.3	7204	14	AF444003	AF444003 Hepatitis E
36	1076	98.3	7207	14	HPECG	D11092 Hepatitis E
37	1076	98.3	7207	14	HPESVP	M73218 Hepatitis E
38	1076	98.3	7212	14	AY230202	AY230202 Hepatitis E
39	1072	97.9	1983	14	AF124407	AF124407 Hepatitis E
40	1072	97.9	2104	14	AF051351	AF051351 Hepatitis E
41	1072	97.9	7193	14	HPEGENA	L25547 Hepatitis E
42	1072	97.9	7202	14	HEVCG	X98292 Hepatitis E
43	1072	97.9	7206	14	AF459438	AF459438 Hepatitis E
44	1072	97.9	7221	14	HPEORFS	L25595 Hepatitis E
45	1071	97.8	2106	14	AF051352	AF051352 Hepatitis E

ALIGNMENTS

RESULT 1	AR278919	Sequence 17 from patent	803 bp	DNA	linear	PAT 10-APR-2003
LOCUS	AR278919	Sequence 17 from patent	US 6514690.			
DEFINITION	AR278919	Sequence 17 from patent	US 6514690.			
ACCESSION	AR278919	Sequence 17 from patent	US 6514690.			
VERSION	AR278919.1	Gi:29713553				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	1 (bases 1 to 803)					
AUTHORS	Li, F., Hui, Z., Anderson, D.A., Logarini, S.A. and Torresi, J.					
TITLE	Immunoreactive antigens of Hepatitis E Virus					
JOURNAL	Patent: US 6514690-A 17 04-FEB-2003;					
FEATURES	source	Location/Qualifiers				
	1..803					
	/organism="unknown"					
	/mol_type="genomic DNA"					
ORIGIN						

Alignment Scores:					
Pred. No.:	1.79e-91	Length:	803		
Score:	1076.00	Matches:	210		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	98.26%	Indels:	0		
DB:	6	Gaps:	0		

US-10-089-292A-2 (1-213) x AR278919 (1-803)

QY 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValIysLeuTyr 20  
|||  
DB 1 CAGCTGTTCTACTCTCGTCCGCTGCTCAGCAATGGCAGCCGACTGTTAAGCTTTAT 60  
|||  
QY 21 ThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 40  
|||  
DB 61 ACATCTGTAGAGATGCTCAGCAGATAGGGTATTGCATCCCGCATGACATGACCTC 120  
|||  
QY 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60  
|||  
DB 121 GGGAGTCTCGTGTAGTTATTGAGTATTATGATTAACCAACATGACGAGCCGACGACA 180  
|||  
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrLeu 80  
|||  
DB 181 CCTTCCCGAGCCCATCGCGCCCTTTCTGTCCTCCGAGCTAATGATGTGCTTTGGCTT 240  
|||  
QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
|||  
DB 241 TCTCTCAGCGCTGCGAGTATGACAGTCCACTTACGGCTCTTCGACCGGCCAGTCTAT 300  
|||  
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
|||  
DB 301 GTCTCTGACTCTGTGACCTTGGTTAATGTTGGACCGCGCGCAGCGCGTTGCCCGGTCA 360  
|||  
QY 121 LeuAspTyrThrIysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140  
|||  
DB 361 CTCGACTGACCAAGTCTACACTTGTATGTCGCCCTTCACCATCCAGCAGTATTCAC 420  
|||  
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheThrGluAlaGlyThrThr 160  
|||  
DB 421 AGACCTCTTTGCTCTGCGCTCGCGGTAGCTCTCTTTGGAGCGCAGTACTACT 480  
|||  
QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrAlaSerAspGlnLeuValGluAsn 180  
|||  
DB 481 AAAGCCGGGTACCTTATATTAATTAACACCACTGCTAGTGACCAACTGCTCGTTGAGAAT 540  
|||  
QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProVal 200  
|||  
DB 541 GCCGCTGGGCATCGGGTTCCTATTTCCACTTACACCACTAGCTGGTCTGGTCCCGTC 600  
|||

RESULT 2  
AR002164  
LOCUS  
DEFINITION Sequence 7 from patent US 5741490.  
ACCESSION AR002164  
VERSION AR002164.1 GI:3963718  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 984)  
Ryves, G.R., Bradley, D.W., Twu, J.-S., Purdy, M.A., Tam, A.W.,  
Krawczynski, K.Z. and Yarbrough, P.D.  
Hepatitis E virus vaccine and method  
Patent: US 5741490-A 7 21-APR-1998;  
Location/Qualifiers  
1. .984  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Alignment Scores:  
Pred. No.: 2,346-91 Length: 984  
Score: 1076.00 Matches: 210  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.26% Indels: 0

DB: 6 Gaps: 0

US-10-089-292A-2 (1-213) x AR002164 (1-984)

QY 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValIysLeuTyr 20  
|||  
DB 181 CAGCTGTTCTACTCTCGCTCCGCTGCTCAGCAATGGCAGCCGACTGTTAAGTGTAT 240  
|||  
QY 21 ThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 40  
|||  
DB 241 ACATCTGTAGAGATGCTCAGCAGATAGGGTATTGCATCCCGCATGACATGACCTC 300  
|||  
QY 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60  
|||  
DB 301 GGAGATCTCGTGTGTTATTGAGTATTATGATTAACCAACATGACCAAGATCGGCCGACG 360  
|||  
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrLeu 80  
|||  
DB 361 CTTTCTCCAGCCCATCGCGCCCTTTCTGTCCTCCGAGCTAATGATGTGCTTTGGCTC 420  
|||  
QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
|||  
DB 421 TCTCTCAGCGCTGCGAGTATGACAGTCCACTTATGCTCTTCGACTGGCCAGTATTAT 480  
|||  
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
|||  
DB 481 GTTCTGACTCTGTGACCTTGGTTAATGTTGGACCGCGCGCAGCGCGTTGCCCGGTCA 540  
|||  
QY 121 LeuAspTyrThrIysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140  
|||  
DB 541 CTCGATTCGACCAAGTCTACACTTGTATGTCGCCCTTCACCATCCAGCAGTACTCG 600  
|||  
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheThrGluAlaGlyThrThr 160  
|||  
DB 601 AAGACCTCTTTGCTCTGCGCTCGCGGTAGCTCTCTTTGGAGCGCAGCACAAC 660  
|||  
QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrAlaSerAspGlnLeuValGluAsn 180  
|||  
DB 661 AAAGCCGGGTACCTTATATTAATTAACACCACTGCTAGTGACCAACTGCTGTCGAGAAT 720  
|||  
QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProVal 200  
|||  
DB 721 GCCGCTGGGCATCGGGTTCCTATTTCCACTTACACCACTAGCTGGTCTGGTCCCGTC 780  
|||

RESULT 3  
AR170408  
LOCUS  
DEFINITION Sequence 5 from patent US 6291641.  
ACCESSION AR170408  
VERSION AR170408.1 GI:17908367  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 984)  
Fuerst, T.R., McAtee, C., Patrick, J., Yarbrough, P.O. and Zhang, Y.-F.  
Hepatitis E virus antigens and uses therefor  
Patent: US 6291641-A 5 18-SEP-2001;  
Location/Qualifiers  
1. .984  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Alignment Scores:  
Pred. No.: 2,346-91 Length: 984  
Score: 1076.00 Matches: 210  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

Query Match: 98.26% Indels: 0  
DB: 6 Gaps: 0

US-10-089-292A-2 (1-213) x AR170408 (1-984)

Qy 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValIleLeuTyr 20  
Db 181 CAGCTGTTCTACTCCCGTCCGTTGTCTCAGCCAATGGCGAGCGACTGTTAAGTTGTAT 240

Qy 21 ThrSerValGluAsnAlaGlnAspGlyGlyIleAlaIleProHisAspIleAspLeu 40  
Db 241 ACATCTGTAGAGATGCTCAGCAGATAAGGTTATTCGAATCCCGCATGACATTCGACCTC 300

Qy 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60  
Db 301 GGAGAATCTCGTGTGGTTATTTCAGGATTAATGAACCAACATGAACAGATCGGCCGACG 360

Qy 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrLeu 80  
Db 361 CCTTCTCCAGCCCCATCGGCCCTTCTCTGTCTTTCGAGTAATGATGCTTTGGCTC 420

Qy 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
Db 421 TCTCTCACCGTCCGAGTATGACAGTCCACTTATGCTCTTCGACTGGCCAGTTTAT 480

Qy 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
Db 481 GTTCTTGACTCTGTGACCTTGGTTAAATGTTGCGACCGCGCGAGCGGCTTGGCCGGTCG 540

Qy 121 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140  
Db 541 CTCGATTGGACCAAGGTACACTTGACGGTCCGCCCTCTCCACCATCCAGCAGTACTCG 600

Qy 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160  
Db 601 AAGACCTCTTGTGCTCCGCTCCGCGTAAAGTCTCTTTCTGGAGGCGAGGCACAACT 660

Qy 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsn 180  
Db 661 AAAGCCGGGTACCTTATTAATTAATTAACACCACTGCTAGCAGCAACTGCTTGTGAGAA 720

Qy 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200  
Db 721 GCCCGCGGCAACCGGTCGCTATTTCACATTACACCACTAGCCTGGGTGGTCCCGTC 780

Qy 201 SerIleSerAlaValAlaValLeuAlaPro 210  
Db 781 TCCATTCTCGGGTTCGGTTTGGCCCTTACGCCCTC 810

RESULT 4  
LOCUS I73145 984 bp DNA linear PAT 03-APR-1998  
DEFINITION Sequence 7 from patent US568239  
ACCESSION I73145  
VERSION I73145.1 GI:3009284  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 984)  
AUTHORS Reyes,G.R., Tam,A.W. and Yarbough,P.O.  
TITLE Hepatitis E virus peptides and methods  
JOURNAL Patent: US 568239-A 7 11-NOV-1997;  
FEATURES  
source  
Location/Qualifiers  
1. 984  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Alignment Scores: 2.34e-91 Length: 984  
Pred. No.: 1076.00 Matches: 210  
Score: 100.00% Conservative: 0  
Percent Similarity:

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.26% Indels: 0  
DB: 6 Gaps: 0

US-10-089-292A-2 (1-213) x I73145 (1-984)

Qy 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValIleLeuTyr 20  
Db 181 CAGCTGTTCTACTCCCGTCCGTTGTCTCAGCCAATGGCGAGCGACTGTTAAGTTGTAT 240

Qy 21 ThrSerValGluAsnAlaGlnAspGlyGlyIleAlaIleProHisAspIleAspLeu 40  
Db 241 ACATCTGTAGAGATGCTCAGCAGATAAGGTTATTCGAATCCCGCATGACATTCGACCTC 300

Qy 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60  
Db 301 GGAGAATCTCGTGTGGTTATTTCAGGATTAATGAACCAACATGAACAGATCGGCCGACG 360

Qy 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrLeu 80  
Db 361 CCTTCTCCAGCCCCATCGGCCCTTCTCTGTCTTTCGAGTAATGATGCTTTGGCTC 420

Qy 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
Db 421 TCTCTCACCGTCCGAGTATGACAGTCCACTTATGCTCTTCGACTGGCCAGTTTAT 480

Qy 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
Db 481 GTTCTTGACTCTGTGACCTTGGTTAAATGTTGCGACCGCGCGAGCGGCTTGGCCGGTCG 540

Qy 121 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140  
Db 541 CTCGATTGGACCAAGGTACACTTGACGGTCCGCCCTCTCCACCATCCAGCAGTACTCG 600

Qy 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160  
Db 601 AAGACCTCTTGTGCTCCGCTCCGCGTAAAGTCTCTTTCTGGAGGCGAGGCACAACT 660

Qy 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsn 180  
Db 661 AAAGCCGGGTACCTTATTAATTAATTAACACCACTGCTAGCAGCAACTGCTTGTGAGAA 720

Qy 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200  
Db 721 GCCCGCGGCAACCGGTCGCTATTTCACATTACACCACTAGCCTGGGTGGTCCCGTC 780

Qy 201 SerIleSerAlaValAlaValLeuAlaPro 210  
Db 781 TCCATTCTCGGGTTCGGTTTGGCCCTTACGCCCTC 810

RESULT 5  
LOCUS AR232469 984 bp mRNA linear PAT 20-DEC-2002  
DEFINITION Sequence 7 from patent-US-6455492  
ACCESSION AR232469  
VERSION AR232469.1 GI:27274538  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 984)  
AUTHORS Reyes,G.R., Bradley,D.W., Twu,J.-S., Purdy,M.A., Tam,A.W.,  
Krawczynski,K.Z. and Yarbough,P.O.  
TITLE Hepatitis E virus vaccine and method  
JOURNAL Patent: US 6455492-A 7 24-SEP-2002;  
FEATURES  
source  
Location/Qualifiers  
1. 984  
/organism="unknown"  
/mol\_type="mRNA"

ORIGIN  
Alignment Scores: 2.34e-91 Length: 984  
Pred. No.: 210  
Score: 100.00% Conservative: 0  
Percent Similarity:

Score:	1076.00	Matches:	210
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	98.26%	Indels:	0
DB:	6	Gaps:	0
US-10-089-292A-2 (1-213) x AR232469 (1-984)			
QY	1	GlnLeuPheTyrSerArgProValValSerAlaAlaGlnGlyProThrValIysLeuTyr	20
DB	181	CAGCTGTTCTACTCCGCTCCGCTTCTCAGCAATGCGGAGCGACTGTTAAGTTGTAT	240
QY	21	ThrSerValGluAsnAlaGlnGlnAspIysGlyIleAlaIleProHisAspIleAspLeu	40
DB	241	ACATCTGTAGAGATGCTCAGCAGATAGAGTATTGCAATCCCGCATGACATTGACCTC	300
QY	41	GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr	60
DB	301	GGAGAATCTCGTGTGTTATTTCAGGATTATGATAACCAACATGAACAGATCGGCGGACG	360
QY	61	ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu	80
DB	361	CCCTCTCCAGCCCATCGCGCTTCTCTGCTTCCTGAGCTAATGATGTGCTTGGCTC	420
QY	81	SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr	100
DB	421	TCTCTCACCGCTGCCGAGTATGACCACTTATGCTCTTCGACTGGCCAGTTTAT	480
QY	101	ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer	120
DB	481	GTTTCTGACTCTGTGACCTTGTGTTAATGTTGGACCGCGCGAGCGCTTGGCCGGTGC	540
QY	121	LeuAspTyrThrIysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer	140
DB	541	CTCGATTGACCAAGTCACTTTCGCGTCCGCCCTCTCCACCATCCAGGACTCG	600
QY	141	LysThrPhePheValLeuProLeuArgGlyIysLeuSerPheTyrProGluAlaGlyThr	160
DB	601	AAGACCTTCTTGTCTCGCGTCCGCGTAAAGTCTCTTCTGGGAGCGGACCAACT	660
QY	161	LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn	180
DB	661	AAAGCCGGGTACCTTATATATATACACCACTGTAGCGACCACTGCTTGTGCGAAT	720
QY	181	AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal	200
DB	721	GCGCCGGGACCGGTCGCTATTCACCTTACACCACTAGCTGGGTGCTGTCGCCGTC	780
QY	201	SerIleSerAlaValAlaValLeuAlaPro	210
DB	781	TCCATTCTGCGGTGCGGTTTTAGCCCCC	810
RESULT 6			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Partial Sequence of HEV Isolates from North Africa and Pakistan:			
Comparison with Known HEV Sequences			
Unpublished			
REFERENCE			
2 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			
Direct Submission			
TITLE			
HEU40044			
LOCUS			
DEFINITION			
Hepatitis E virus structural protein 2 (orf2) mRNA, partial cds.			
U40044			
VERSION			
U40044.1 GI:4096791			
KEYWORDS			
Hepatitis E virus			
SOURCE			
ORGANISM			
Hepatitis E virus			
REFERENCE			
1 (bases 1 to 1117)			
Van Cuyck-Gandre, H., Clements, N.J., Zhang, H.Y., Caudill, J.C.,			
Cohen, S.G., Coursaget, P., Buisson, Y., Warren, R.L. and Longer, C.F.			

```
QY 201 SerIleSerAlaValAlaValLeuAlaPro 210
Db 832 TCCATTTCGCGGTGCTGTTTAGCCCCC 861

RESULT 7
LOCUS AR002166 1311 bp linear PAT 04-DEC-1998
DEFINITION Sequence 9 from patent US 5741490-A
ACCESSION AR002166
VERSION AR002166.1 GI:3963720
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1311)
AUTHORS Reyes,G.R., Bradley,D.W., Twu,J.-S., Purdy,M.A., Tam,A.W.,
Krawczynski,K.Z. and Yarbough,P.D.
TITLE Hepatitis E virus vaccine and method
JOURNAL Patent: US 5741490-A 9 21-APR-1998;
FEATURES
LOCATION/Qualifiers
SOURCE 1. 1311
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 3.4e-91 Length: 1311
Score: 1076.00 Matches: 210
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.26% Indels: 0
DB: 6 Gaps: 0

US-10-089-292A-2 (1-213) x AR002166 (1-1311)
QY 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 20
Db 508 CAGCTGTTCTACTCCCGTCCCGTGTCTCAGCAATGCGGAGCGGACTGTTAAGTTGTAT 567
QY 21 ThrSerValGluAsnAlaGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 40
Db 568 ACACTGTAGAGATGCTCAGCAGATAGAGGTATTCACCAATCCCGCATGACATTGACCTC 627
QY 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60
Db 628 GGAGAACTCTGTGTTATTCAGGATATGATACCAACATGACCAAGATCGGCGGACG 687
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80
Db 688 CTTTCTCCAGCCCCATCGCGCCCTTCTCTGCTCTCGAGCTAATGATGTGCTTTGGCTC 747
QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100
Db 748 TCTCTCACCGTTCGCGAGTATGACGATGATGCTTATGCTCTTCGACTGGCCGAGTTAT 807
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120
Db 808 GTTCTGACTCTGTGACCTTGGTTAATGTCGACCGCGCGAGCGGCTTCCCGGTCG 867
QY 121 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140
Db 868 CTCGATTGGACCAAGTCCACACTTGACGTCGCCCTCTCCACCATCCAGCAGTACTCG 927
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160
Db 928 AAGACCTTCTTTGTCTCTCCGCTCCGCGTAAAGTCTCTCTTTCTGGAGGCGAGGCAACT 987
QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180
Db 988 AAAGCGGGTACCTTTATATTAATTAACCACTGCTAGGACCAACTGCTGTCGAGAAAT 1047
QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProVal 200

Db 1048 GCCCCCGGCGACCGGTCGCTATTTCACCTTACACCTAGCTGGGTCGCTGCTCCCGTC 1107
QY 201 SerIleSerAlaValAlaValLeuAlaPro 210
Db 1108 TCCATTTCGCGGTGCGGTTTAGCCCCC 1137

RESULT 8
LOCUS AR232471 1311 bp mRNA linear PAT 20-DEC-2002
DEFINITION Sequence 9 from patent US 6455492.
ACCESSION AR232471
VERSION AR232471.1 GI:27274540
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1311)
AUTHORS Reyes,G.R., Bradley,D.W., Twu,J.-S., Purdy,M.A., Tam,A.W.,
Krawczynski,K.Z. and Yarbough,P.O.
TITLE Hepatitis E virus vaccine and method
JOURNAL Patent: US 6455492-A 9 24-SEP-2002;
FEATURES
LOCATION/Qualifiers
SOURCE 1. 1311
/organism="unknown"
/mol_type="mRNA"
```

Db	1327	AAAGCCGGGTACCC	TATAATATATATATACACCACTGCTACGACCACTGCTTGTGCGAAT	1385	
Qy	181	AlaAlaGlyHisArgVal	lalleSerThrTyThrSerLeuGlyAlaGlyProVal	200	
Db	1387	GCCCGCGGGCACC	GGTGCATTTCCACTTACACCACTAGCCTGGGTGCTGGTCCCGTC	1446	
Qy	201	SerIleSerAlaVal	AlaValLeuAlaPro	210	
Db	1447	TCCATTCTCGCGT	TTCGGTTTAGCCCCC	1476	
RESULT 10					
LOCUS	E17105	1647 bp	DNA	linear	PAT 28-JUL-1999
DEFINITION	EA encoding HEV hollow particle.				
ACCESSION	E17105				
VERSION	E17105.1	GI:5711788			
KEYWORDS	JP 1998234383-A/1.				
SOURCE	unidentified				
ORGANISM	unclassified,				
REFERENCE	1 (bases 1 to 1647)				
AUTHORS	Takeda,N., Ri,T. and Miyamura,T.				
TITLE	HEPATITIS E VIRUS HOLLOW PARTICLE, GENE ENCODING THE SAME AND PREPARATION OF RECOMBINATION VECTOR CONTAINING THE GENE AND UTILIZATION OF THE RECOMBINATION VECTOR				
JOURNAL	Patent: JP 1998234383-A 1 08-SEP-1998;				
COMMENT	DENKA SEIKEN CO LTD, KOKURITSU KANSENSHIYOU KENKYUSHO OS Hepatitis E virus PN JP 1998234383-A/1 PD 08-SEP-1998 PF 28-FEB-1997 JP 1997062445 PI TAKEDA NAOKAZU, RI TENSEI, MIYAMURA TATSUO PC C12N15/09,C07K14/08,C12N5/10,C12P21/02; CC strandedness: Double; CC topology: Linear; FH Key Location/Qualifiers FH FT source 1..1647 FT /organism='Hepatitis E virus' FT mat_peptide FT 1..1647 /product='hollow particle'. FT Location/Qualifiers 1..1647 /organism='unidentified' /mol_type='genomic DNA' /db_xref='taxon:32644'				
FEATURES	source				
ORIGIN					
Alignment Scores:					
Pred. No.:	4.59e-91	Length:	1647		
Score:	1076.00	Matches:	210		
Percent Similarity:	100.00%	Conservative:	0		
Best Local Similarity:	100.00%	Mismatches:	0		
Query Match:	98.26%	Indels:	0		
DB:	6	Gaps:	0		
US-10-089-292A-2 (1-213) x E17105 (1-1647)					
Qy	1	GlnLeuPheTyrSerArgProVal	SerAlaAsnGlyGluProThrVal	LeuTyr	20
Db	847	CAGCTGTTCTACTCCCGTCCCGTTGTCTCGSCCAATGCGGACGAC	CTGTTAAAGTTGAT	906	
Qy	21	ThrSerValGluAsnAlaGlnGlnAspIysGlyIle	alalleProHisAspIleAspLeu	40	
Db	907	ACATCTGTAGAGAAATGCTCAGCAGATAAGGGTATTGCAATCCCGCATGACATTGACCTC	966		
Qy	41	GlyGluSerArgValValIleGlnAspTyAspAsnGlnHisGluGlnAspArgProThr	60		
Db	967	GGAGAACTCGTGTGGTATTTCAGGATTATGACACCAACCATGAACAGATCGGCCGACG	1026		
Qy	61	ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTripleu	80		
Db	1027	CTTTCTCAGCCCATCGGCCCTTCTCTGTCTCGAGTAATGATGTGCTTTGGCTC	1086		

QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
Db 1087 TCTCTCACCGCTGCCAGTATGACCACTTATGCTCTTCGACTGGCCCACTTTAT 1146  
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
Db 1147 GTTCTCTGACTCTGTGACCTTGGTTAATGTTGGACCGCGCGAGCGCTTGGCCGGTCG 1206  
QY 121 LeuAspTprThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140  
Db 1207 CTCGATTTGACCAAGTCCACTTGACGTCGCCCTCTCCACCATCCAGCAGTACTCG 1266  
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160  
Db 1267 AAGACCTCTCTTGTCTCGCTCGCGGTAAAGCTCTCTTCTTGGGAGCGGACCAACT 1326  
QY 161 LysAlaGlyTyrProTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180  
Db 1327 AAAGCCGGGTACCCCTTAATTAATTAACACCACTGCTAGCACCACCACTGCTGCGAGAAT 1386  
QY 181 AlaAlaGlyHisArgValAlaAlaIleSerThrTyrThrSerLeuGlyAlaGlyProVal 200  
Db 1387 GCCGCGGGGACCGGCTGCTATTTCACCTTACACCACTAGCCTGGGTGCTGGTCCCGTC 1446  
QY 201 SerIleSerAlaValAlaValLeuAlaPro 210  
Db 1447 TCCATCTCTCGCGTGTGCTTTTAGCCCCC 1476

RESULT 11  
AR278918  
LOCUS AR278918 1658-bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 15 from patent-US-6514690.  
ACCESSION AR278918  
VERSION AR278918.1 GI:29713552  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1658)  
AUTHORS Li, F., Hui, Z., Anderson, D.A., Logarini, S.A. and Torresi, J.  
TITLE Immunoreactive antigens of Hepatitis E Virus  
JOURNAL Patent: US 6514690-A 15 04-FEB-2003;  
FEATURES Location/Qualifiers  
source 1..1658  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores:  
Pred. No.: 4.63e-91 Length: 1658  
Score: 1076.00 Matches: 210  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.26% Indels: 0  
DB: 6 Gaps: 0  
US-10-089-292A-2 (1-213) x AR278918 (1-1658)

QY 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 20  
Db 884 CAGCTGTTCTACTCTCTCGTCCCGTCTCTCAGCAATGGCGAGCCGACTGTTAAGCTTTAT 943  
QY 21 ThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspLeu 40  
Db 944 ACATCTGATAGAAATGCTCAGCAGGATAAGGTATTGCAATCCCGATGACATCAGCTTC 1003  
QY 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60  
Db 1004 GGGAGTCTCGTGTAGTATTTCAGATTATGACACCAACATGACGAGCCGACCGACA 1063  
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80

Db 1064 CTTCCCGAGCCCATCGCGCCCTTTTCTGCTCTCCGAGCTAATGATGTGCTTGGCTT 1123  
QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
Db 1124 TCTCTCACCGCTGCCAGTATGACCACTTACGCTCTTCGACCGGCCAGCTTAT 1183  
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
Db 1184 GTCTCTGACTCTGTGACCTTGGTTAATGTTGGACCGCGCGAGCGCTTGGCCGGTCA 1243  
QY 121 LeuAspTprThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140  
Db 1244 CTCGATTTGACCAAGTCCACTTGATGTCGCCCTCTCCACCATCCAGCAGTATTCA 1303  
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160  
Db 1304 AAGACCTCTCTTGTCTCGCTCGCGGTAAAGCTCTCTTCTTGGGAGCGAGTACTACT 1363  
QY 161 LysAlaGlyTyrProTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180  
Db 1364 AAAGCCGGGTACCCCTTAATTAATTAACACCACTGCTAGTACCACTGCTGCTGAGAAT 1423  
QY 181 AlaAlaGlyHisArgValAlaAlaIleSerThrTyrThrSerLeuGlyAlaGlyProVal 200  
Db 1424 GCCGCTGGGACATCGGTTGCTATTTCACCTTACACCACTAGCCTGGGTGCTGGTCCCGTC 1483  
QY 201 SerIleSerAlaValAlaValLeuAlaPro 210  
Db 1484 TCTATTTCGCGGTGCTGTTTATAGCCCCC 1513

RESULT 12  
AR278920  
LOCUS AR278920 1686-bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 19 from patent US 6514690.  
ACCESSION AR278920  
VERSION AR278920.1 GI:29713554  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1686)  
AUTHORS Li, F., Hui, Z., Anderson, D.A., Logarini, S.A. and Torresi, J.  
TITLE Immunoreactive antigens of Hepatitis E Virus  
JOURNAL Patent: US 6514690-A 19 04-FEB-2003;  
FEATURES Location/Qualifiers  
source 1..1686  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores:  
Pred. No.: 4.73e-91 Length: 1686  
Score: 1076.00 Matches: 210  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.26% Indels: 0  
DB: 6 Gaps: 0  
US-10-089-292A-2 (1-213) x AR278920 (1-1686)

QY 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 20  
Db 884 CAGCTGTTCTACTCTCTCGTCCCGTCTCTCAGCAATGGCGAGCCGACTGTTAAGCTTTAT 943  
QY 21 ThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspLeu 40  
Db 944 ACATCTGATAGAAATGCTCAGCAGGATAAGGTATTGCAATCCCGATGACATCAGCTTC 1003  
QY 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60  
Db 1004 GGGAGTCTCGTGTAGTATTTCAGATTATGACACCAACATGACGAGCCGACCGACA 1063  
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80

```
Db 1064 CTTTCCCGAGCCCATCGGCCCTTTCTGCTCCGAGTATGATGCTTTGGCTT 1123
Qy 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyGlySerSerThrGlyProValTyr 100
Db 1124 TCTCTCACCGCTGCGAGTATGACAGTCCACTTACGGCTCTTCGACCGGCCAGTCTAT 1183
Qy 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120
Db 1184 GTCTCTGACTCTGTGACCTTGTTAATGTGCGACCGCGCGAGCGCTTCCCGGTCA 1243
Qy 121 LeuAspTyrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140
Db 1244 CTCGACTGGACCAAGTACACTTGATGGTCGCCCTTCCACCATCCAGCAGTATCA 1303
Qy 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThr 160
Db 1304 AAGACCTTCTTGTCTGCGCTCGCGGTAAAGTCTCTCTTTTGGAGGAGTACTACT 1363
Qy 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180
Db 1364 AAAGCGGGTACCTTATAATTAATTAACCACTGCTAGTACCACTGCTGTTGAGAT 1423
Qy 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200
Db 1424 GCGCTGGGCATCGGGTGTCTATTTCCACTTACACCACTAGCTGGTGGTCCCGTC 1483
Qy 201 SerIleSerAlaValAlaValLeuAlaPro 210
Db 1484 TCTATTCCGCGTGTGCTGTTTGTAGCCCC 1513

RESULT 13
HEVBCN6 1946 bp RNA linear VPL 17-NOV-1998
LOCUS Hepatitis E virus nonstructural polyprotein and structural protein
DEFINITION genes, partial cds; and unknown gene.
ACCESSION AF058684
VERSION AF058684.1 GI:3882959
KEYWORDS
SEGMENT
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1946)
Pina,S., Jofre,J., Emerson,S.U., Purcell,R.H. and Girones,R.
Characterization of a strain of infectious hepatitis E virus
isolated from sewage in an area where hepatitis E is not endemic
Appl. Environ. Microbiol. 64 (11), 4485-4488 (1998)
99797311
PUBMED
REFERENCE
2 (bases 1 to 1946)
Pina,S., Jofre,J., Shapiro,M., Emerson,S.U., Purcell,R.H. and
Girones,R.
Direct Submission
Submitted (09-APR-1998) Department of Microbiology, School of
Biology, University of Barcelona, Diagonal Avenue 645, Barcelona
08026, Spain
Location/Qualifiers
1. .1946
/organism="Hepatitis E virus"
/mol_type="genomic RNA"
/strain="BCN"
/db_xref="taxon.12461"
/country="Spain:Barcelona"
<1. .95
/notes="ORF1"
/codon_start=3
/product="nonstructural polyprotein"
/protein_id="AAC77805.1"
/db_xref="GI.3882967"
/translation="LQAVADGKAHFTESVKPVLDTNSILCRVE"
misc_feature <1. .92
```

```
CDS
/notes="Region: putative RNA-directed RNA polymerase"
92. .463
/notes="ORF3"
/codon_start=1
/product="unknown"
/protein_id="AAC77807.1"
/db_xref="GI:3882969"
/translations="MNNMSPAAPMGSRPCALGIFCCSSCFCLCCPRRHVRVSLAAVV
GGAAPVAVVGVGTGLILRPSQPIQTPTSPSPMLRGLDLVFNPPDHPAPLGV
TRSAAPLPHVVDLPQLGPRR"
133. >1946
/notes="ORF2"
/codon_start=1
/product="structural protein"
/protein_id="AAC77808.1"
/db_xref="GI:3882970"
/translation="MRPRPILLFLMPLPMLPAPPQSSRRRRSGSGSGGFGWD
RVSDPPAIPIIHTNPPADVTAAAGAGRVQPARPLASAWRDOAPAAASRRRP
TTTCAAPLATAFAHDTFPVDSRQAILRRQYNLSTPSTSVATGTLNVLVAAFL
SPLLPLQDGTNTHIMATEASYAQYVARATIRYPLVPNAVGGYALISFWPOTTTT
PTSDMNSITSDVRIIVQPGIASLVIPISERLHYRQRRSVETSGAQEEATSGLV
MLCTHSGPVSNSYNTPTVYTGALGLDPALELRNLTPGNTNTRVRSSTARHLRRG
ADGTAELTTTAATREMKDLYFTSNNGVGEIGRGLALFLNLADTLGLGLPTELISSAG
GQIFYSPVVSANGEPVTKLYTSVENAQDQKIALPHDIDLGESRVVIQYDNDQHEOD
RPTSPAPSRPFSVLRANDVLSLTAAYIDQSYTSGSTGTPVVSVDSTLVNATGAQ
AVARSLDWTKTVDLRPLSTIQOYSKTFVFLRLGKLSFWFAGITTKAGYPTNYNTIAS
DQLLVENAAGRHVAISTYTTSLGAGPVVISAVAVLAAPS"
133. >1946
misc_feature
/notes="Region: putative capsid protein"
ORIGIN
Alignment Scores:
Pred. No.: 5,7e-91 Length: 1946
Score: 1076.00 Matches: 216
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.26% Indels: 0
DB: 14 Gaps: 0
US-10-089-292A-2-(1-213) x HEVBCN6 (1-1946)
Qy 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 20
Db 1312 CAGCTGTTCTACTCCGTCCTGCTCAGCCAAATGGCGAGCGACTGTTAAGCTGAT 1371
Qy 21 ThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 40
Db 1372 ACATCTGTAGAGATGCTCAGCAGGATAAGGATATCGCAATCCCGCATGATATTGACCTC 1431
Qy 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60
Db 1432 GGAGAGTCTCGTGTGTTTATTTCAGGATATTATATCAACCAATGATGAACAGCCGCGCAGC 1491
Qy 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTyrLeu 80
Db 1492 CCTTCTCCAGCCCCCTCGGCCCTTCTCTGTTCTTCGAGTAAATGATGTGCTTTGGCTC 1551
Qy 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100
Db 1552 TCTCTCACCGCTGCGAGTATGACCACTATGCTTCTTCGACTGGCCAGCTTAT 1611
Qy 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120
Db 1612 GTTCTGACTCTGTGACCTTGTTAATGTGCGACCGCGCGCGAGCGCTTCCCGGTCCG 1671
Qy 121 LeuAspTyrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140
Db 1672 CTCGACTGGACCAAGTACACTTTCAGCGTCCCTCTCTCCACCATCCAGCAGTACTCG 1731
Qy 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThr 160
Db 1732 AAGACCTTCTTGTCTGCGCTCGCGGCAAGCTCTCTTTTGGAGGCGGCGCAACT 1791
```



```

Qy 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValIcuAsn 180
Db 1792 AAGCCGGGTACCCCTTAATTAATACACACCACTGCTAGCGACCAACTCTGTCGAGAA 1851
Qy 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrSerLeuGluAlaGlyProVal 200
Db 1852 GCCCGCGGACCGGGTGGCATCTCCACTTACACCACTAGCTGGGTGCTGGCCCGTC 1911
Qy 201 SerIleSerAlaValAlaValLeuAlaPro 210
Db 1912 TCCATTTCTCGCGGTGCGCTCTAGCCCCC 1941

RESULT 14
LOCUS Hepatitis E virus ORF2. 1983 bp DNA linear PAT 28-JUL-1999
DEFINITION Hepatitis E virus ORF2.
ACCESSION E17107
VERSION E17107.1 GI:5711790
KEYWORDS JP 1998234383-A/3.
SOURCE unidentifed
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1983)
AUTHORS Takeda,N., Ri,T. and Miyamura,T.
TITLE HEPATITIS E VIRUS HOLLOW PARTICLE, GENE ENCODING THE SAME AND
PREPARATION OF RECOMBINATION VECTOR CONTAINING THE GENE AND
UTILIZATION OF THE RECOMBINATION VECTOR
JOURNAL Patent: JP 1998234383-A 3 08-SEP-1998;
DENKA SEIKEN CO LTD, KOKURITSU KANSENSHIYOU KENKYUSHO
COMMENT OS Hepatitis E virus
PN JP 1998234383-A/3
PD 08-SEP-1998
PF 28-FEB-1997 JP 1997062445
PI TAKEDA NAKAZU, RI TENSEI, MIYAMURA TATSUO
PC C12N15/09, C07K14/08, C12N5/10, C12P21/02;
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1. .1983 /organism='Hepatitis E virus' FT CDS
FT 1. .1983 /product='structural protein'.
FEATURES
source Location/Qualifiers
1. .1983 /organism='unidentifed'
/mol_type='genomic DNA'
/db_xref='taxon:32644'

ORIGIN
Alignment Scores: 5,85e-91 Length: 1983
Pred. No.: 1076.00 Matches: 210
Score: 1076.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 98.26% Gaps: 0
DB: 6

US-10-089-292A-2 (1-213) x E17107 (1-1983)
Qy 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValIysLeuTyr 20
Db 1180 CAGCTGTTCTACTCCGCTCCCGTGTCTCGGCCAATGGCGAGCCGACTGTTAAGTTGTAT 1239
Qy 21 ThrSerValGluAsnAlaGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 40
Db 1240 ACATCTGTAGAGATGCTCAGCAGGATAAGGTTATTGCAATCCCGCATGACATTCACCTC 1299
Qy 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60
Db 1300 GGAGAATCTCGTGTGTTATTTCAGGATTATGACACCAACATGAACATGACGCGGACG 1359
Qy 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80

```

```

Db 1360 CCTTCTCAGGCCCATCGGCCCTTTCTCTGCTCTTCGAGCTAATGATGCTTGGGTC 1419
Qy 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100
Db 1420 TCTCTCACCCTGCGGAGTATGACCACTTATGCTCTTCGACTGGCCAGTTTAT 1479
Qy 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120
Db 1480 GTTTCTGACTCTGTGACCTTGTGTTAATGTTGCGACCGCGCGAGCGGTTGCCCGGTCG 1539
Qy 121 LeuAspTrpThrLysValThrLeuAspGlyValArgProLeuSerThrIleGlnGlnTyrSer 140
Db 1540 CTCGATTGGACCAAGTTCACACTTGACGTCGCCCTCTCCCATCCAGCAGTACTCG 1599
Qy 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160
Db 1600 AAGACCTTCTTTGTCTCGCGCTCCGCGGTAAGCTCTCTCTTCTGGAGGAGGCACAACT 1659
Qy 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180
Db 1660 AAGCCGGGTACCCCTTATATAACACCACTGCTAGCGACCAACTCTCTTGTGCGAGAT 1719
Qy 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200
Db 1720 GCCCGCGGACCGCGTGCCTATTTCACCTTACACCACTAGCTGGGTGCTGGTCCCCGTC 1779
Qy 201 SerIleSerAlaValAlaValLeuAlaPro 210
Db 1780 TCCATCTCTCGCGTGTGCTGTTTATAGCCCCC 1809

RESULT 15
LOCUS I73139 1983 bp DNA linear PAT 03-APR-1998
DEFINITION Sequence 1 from patent:US-5686239.
ACCESSION I73139
VERSION I73139.1 GI:3009278
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1983)
AUTHORS Reyes,G.R., Tam,A.W. and Varbough,P.O.
TITLE Hepatitis E virus peptides and methods
JOURNAL Patent: US 5686239-A 1 11-NOV-1997;
FEATURES Location/Qualifiers
source 1. .1983 /organism='unknown'
/mol_type='unassigned DNA'

ORIGIN
Alignment Scores: 5,85e-91 Length: 1983
Pred. No.: 1076.00 Matches: 210
Score: 1076.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 98.26% Gaps: 0
DB: 6

US-10-089-292A-2 (1-213) x I73139 (1-1983)
Qy 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValIysLeuTyr 20
Db 1180 CAGCTGTTCTACTCCGCTCCCGTGTCTCAGCCAATGGCGAGCCGACTGTTAAGTTGTAT 1239
Qy 21 ThrSerValGluAsnAlaGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 40
Db 1240 ACATCTGTAGAGATGCTCAGCAGGATAAGGTTATTGCAATCCCGCATGACATTCACCTC 1299
Qy 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60
Db 1300 GGAGAATCTCGTGTGTTATTTCAGGATTATGATAACCAACATGAACATGACGCGGACG 1359

```

```
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80
Db |||||
1360 CCTTCTCCAGCCCATCGCGCCTTCTCTGCTTCGAGCTAATGATGTGCTTTGGCTC 1419
QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100
Db |||||
1420 TCTCTCACCGCTGCCGAGTATACCACTTATGGCTCTTCGACTGGCCCAAGTTTAT 1479
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyValAlaGlnAlaValAlaArgSer 120
Db |||||
1480 GTTTCGTGACTCTGTGACCTTGGTTAATGTTGGACCGCGCGCGAGCCGTTGCCCGGTG 1539
QY 121 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140
Db |||||
1540 CTCGATTGGACCAAGGTCACACTTGACGGTCGCCCTCTCCACCATCCAGCAGTACTCG 1599
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160
Db |||||
1600 AAGACCTTCTTTGCTGCTGCCGCTCCGCGGTAAAGCTCTCTTTCTGGAGGCGGCACAACT 1659
QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsn 180
Db |||||
1660 AAAGCGGGTACCTTATATATTAACACCACTGCTAGCAGCAACTGCTTGTTCAGGAT 1719
QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200
Db |||||
1720 GCCGCGGGGACCGGGTCGCTATTTCCACTTACACCACTAGCCTGGGTGCTGGTCCCGTC 1779
QY 201 SerIleSerAlaValAlaValLeuAlaPro 210
Db |||||
1780 TCCATTCTGCGGTTGCCGTTTTAGCCCC 1809
```

Search completed: October 28, 2004, 11:15:54  
Job time : 2721 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: October 28, 2004, 09:42:54 ; Search time 330 Seconds  
(without alignments)  
3388.263 Million cell updates/sec

Title: US-10-089-292A-2

Perfect score: 1095

Sequence: 1 QLFVSRPVSGANGPTVKLY.....SLGAGPVISAVAVLAPPPR 213

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO\_spool/US10089292/runat\_27102004\_153927\_10148/app\_query.fasta\_1.391  
-DB=N\_Geneseq\_23Sep04 -QPMT=fastap -SUFFIX=ring -MINMATCH=0 -1-LOOPCL=0  
-LOOPEXT=0 -UNIT5=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFT=pco -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10089292 @CGN 1 1 352 @runat\_27102004\_153927\_10148 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6  
-FCGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_23Sep04.\*

1: geneseqn1980s.\*  
2: geneseqn1990s.\*  
3: geneseqn2000s.\*  
4: geneseqn2001as.\*  
5: geneseqn2001bs.\*  
6: geneseqn2002as.\*  
7: geneseqn2002bs.\*  
8: geneseqn2003as.\*  
9: geneseqn2003bs.\*  
10: geneseqn2003cs.\*  
11: geneseqn2003ds.\*  
12: geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1095	100.0	810	4 AAS01434	Aas01434 Hepatitis
2	1095	100.0	1989	6 ABL60054	Ab160054 Hepatitis
3	1095	100.0	1989	10 ADB97809	Adb97809 HEV relat
4	1076	98.3	984	2 AAT27111	Aat27111 Hepatitis
5	1076	98.3	984	2 AAT96965	Aat96965 Hepatitis
6	1076	98.3	1647	2 AAT27109	Aat27109 Hepatitis

7	1076	98.3	1647	2 AAV61686	Aav61686 Hepatitis
8	1076	98.3	1983	2 AAT96959	Aat96959 Hepatitis
9	1076	98.3	1983	2 AAV61688	Aav61688 Hepatitis
10	1076	98.3	1984	2 AAQ86592	Aaq86592 Hepatitis
11	1076	98.3	1990	6 ABL60053	Ab160053 Hepatitis
12	1076	98.3	1990	10 ADB97808	Adb97808 HEV ORF2
13	1076	98.3	2049	2 AAT27107	Aat27107 Hepatitis
14	1076	98.3	2054	4 AAS01448	Aas01448 Hepatitis
15	1076	98.3	2094	2 AAQ47129	Aaq47129 HEV thrd
16	1076	98.3	2094	2 AAQ46813	Aaq46813 Burma str
17	1076	98.3	7158	2 AAT27394	Aat27394 Hepatitis
18	1076	98.3	7168	2 AAQ45197	Aaq45197 HEV strai
19	1076	98.3	7168	2 AAV71604	Aav71604 Hepatitis
20	1076	98.3	7194	2 AAV61690	Aav61690 Hepatitis
21	1076	98.3	7195	2 AAV66321	Aav66321 ET-NANB (
22	1076	98.3	7195	2 AAV54729	Aav54729 DNA seque
23	1076	98.3	7195	3 AAA99259	Aaa99259 Hepatitis
24	1076	98.3	7195	4 AAF83495	Aaf83495 ET-NANB v
25	1076	98.3	7195	6 AAL50386	Aal50386 Hepatitis
26	1076	98.3	7204	9 ADA50062	Ada50062 SK-HEV-3
27	1076	98.3	7204	9 ADA50065	Ada50065 Hepatitis
28	1076	98.3	7204	9 ADA50064	Ada50064 Hepatitis
29	1076	98.3	7204	9 ADA50063	Ada50063 SK-HEV-2
30	1073	98.0	7195	2 AAQ14412	Aaq14412 Forward s
31	1067	97.4	7195	10 ADD24374	Add24374 Hepatitis
32	1051	96.0	7099	5 ADH13466	Adh13466 Hepatitis
33	1051	96.0	7194	5 ADH13469	Adh13469 Hepatitis
34	1043	95.3	984	2 AAT27112	Aat27112 Hepatitis
35	1043	95.3	984	2 AAT96966	Aat96966 Hepatitis
36	1043	95.3	1647	2 AAT27110	Aat27110 Hepatitis
37	1043	95.3	1980	2 AAT96960	Aat96960 Hepatitis
38	1043	95.3	2058	2 AAT27108	Aat27108 Hepatitis
39	1043	95.3	2100	2 AAQ47130	Aaq47130 HEV thrd
40	1043	95.3	7171	2 AAQ14413	Aaq14413 Composite
41	1043	95.3	7171	2 AAV66322	Aav66322 ET-NANB (
42	1043	95.3	7171	2 AAV54730	Aav54730 Composite
43	1043	95.3	7171	3 AAA99260	Aaa99260 Hepatitis
44	1043	95.3	7171	4 AAF83496	Aaf83496 Mexican s
45	1043	95.3	7171	6 AAL50387	Aal50387 Hepatitis

#### ALIGNMENTS

RESULT 1

AAS01434

ID AAS01434 standard; cDNA; 810 BP.

XX

AC AAS01434;

XX

DT 18-JUL-2001 (first entry)

XX

XX Hepatitis E virus (HEV) E2 cDNA fragment derived from ORF2.

XX Hepatitis E virus; HEV; E2; PE2; ORF2; immunoreactive viral peptide;

XX antigenic activity; viral infection; mutant; ss.

XX

OS Hepatitis E virus; Chinese strain D11092.

OS Synthetic.

XX

XX Key Location/Qualifiers

FT misc\_feature 1..810

FT /tag= b

FT /note= "Corresponds to residues 1180-1990 of ORF2"

FT CDS 1..642

FT /tag= a

FT /partial

FT /product= "pE2 peptide"

FT /note= "Specifically claimed. This sequence lacks a start

FT codon

FT mutation replace(631..632, CAC)

FT /tag= c

XX WO200122916-A2.

PN

XX 05-APR-2001.  
 XX 28-SEP-2000; 2000WO-1B001393.  
 XX 30-SEP-1999; 99CA-02283538.  
 XX (UYHK-) UNIV HONG KONG.  
 XX Ng MH, Im S, Zhang J;  
 XX WPI; 2001-266109/27.  
 DR P-PSDB; AAU00496.  
 XX  
 PT New immunoreactive peptide pE2, cloned from the genome of a Chinese  
 PT strain of hepatitis E virus (HEV), for detecting and diagnosing HEV  
 PT infection and for use as a vaccine for prevention of the viral infection  
 PT in humans.  
 XX  
 PS Claim 5; Fig 2B-2D; 101pp; English.  
 XX  
 CC The present sequence for hepatitis E virus (HEV) E2 cDNA fragment encodes  
 CC for a novel peptide, pE2 which is cloned from ORF2 (AA501448) of the HEV  
 CC Chinese strain D11092 genome. During the cloning of the E2 fragment a  
 CC single base deletion (presumed to be a PCR amplification error) at  
 CC position 1811 of ORF2 resulted in a frame shift causing the premature  
 CC termination at a new stop codon at position 1820-1822, and giving rise to  
 CC a smaller peptide, pE2. Peptide pE2 is a highly immunoreactive viral  
 CC peptide which can be used in the development of diagnostic methods for  
 CC the detection of HEV. Peptide pE2 which has antigenic activity is also  
 CC useful as a vaccine for immunising an individual against infection from  
 CC HEV and for determining the presence of HEV antibodies in a biological  
 CC test sample such as human blood serum or plasma for detection and  
 CC diagnosis of HEV infection. An antibody against pE2 is useful for  
 CC detecting an HEV particle, in a biological test sample. The use of  
 CC recombinant peptides such as pE2 peptide have advantage over the prior  
 CC art of using live intact virus particles by avoiding the potential risk  
 CC of infectivity. The ORF3 (AA501449) of HEV Chinese strain D11092 is used  
 CC to clone a cDNA fragment E3 (AA5001450) which encodes pE3 peptide  
 XX  
 SQ Sequence 810 BP; 150 A; 249 C; 192 G; 219 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1.09e-103 Length: 810  
 Score: 1095.00 Matches: 213  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 4 Gaps: 0

US-10-089-292A-2 (1-213) x AA501434 (1-810)

QY 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValIysLeuTyr 20  
 DB 1 CAGCTGTTCTACTCTCGTCCGCTCTCAGCAATGCGGAGCCGACTGTTAAGCTTTAT 60  
 QY 21 ThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 40  
 DB 61 ACATCTGTAGAGATGCTCAGCAGATAGAGGTATGCAATCCCGATGACATCCACCTC 120  
 QY 41 GlyGluSerArgValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60  
 DB 121 GGGGAGTCTCGTGTAGTTATTTCAGGATTATGACCAACATGAGCAGGACCGACGACA 180  
 QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80  
 DB 181 CCTTCCCGAGCCCATCGCGCCCTTTCTGTCCTCCGAGCTAATGATGTCTTTGGCTT 240  
 QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
 DB 241 TCTCTCACCCTGCGGAGTATGACCAAGTCCATTCAGGCTCTTCGACCGGCGGCTATAT 300  
 QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120

DB 301 GTCTCTGACTCTGTGACCTTGGTTAATGTTGGACCGCGCGCAGGCCGTTGCCCGGTCA 360  
 QY 121 LeuAspTyrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140  
 DB 361 CTCGACTGGACCAAGGTACACTTGTATGGTCCGCCCTTTTCACCATCCAGCAGTATTCA 420  
 QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThr 160  
 DB 421 AAGACCTTCTTTGCTCTGCGGCTCCGCGTAAGCTCTCTCTTTGGAGGCAGGTACTACT 480  
 QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180  
 DB 481 AAGCGCGGTACCTTATAATTAACACCACTGCTAGTGACCACTGCTCGTTGAGAAT 540  
 QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200  
 DB 541 GCCGCTGGGCATCGGGTTGCTATTTCACCTTACACCACTAGCTGGGTGCTGCTCCGTC 600  
 QY 201 SerIleSerAlaValAlaValLeuAlaProProArg 213  
 DB 601 TCTATTTCGGGTTGCTGTTTACCCCCCTCCGCGC 639  
 RESULT 2  
 ABL60054  
 ID ABL60054 standard; DNA; 1989 BP.  
 XX  
 AC ABL60054;  
 XX  
 DT 16-AUG-2002 (first entry)  
 XX  
 DE Hepatitis E virus ORF-2 DNA #2.  
 XX  
 KW Hepatitis E virus; ORF-2; HEV; virucide; immunisation; gene; ds.  
 XX  
 OS Hepatitis E virus.  
 XX  
 PN WO200240681-A1.  
 XX  
 PD 23-MAY-2002.  
 XX  
 PF 30-SEP-2001; 2001WO-CN001469.  
 XX  
 PR 30-SEP-2000; 2000CN-00130634.  
 XX  
 PA (YANG-) YANG SHENG TANG CO LTD.  
 XX  
 PI Xia N, Zhang J, Li S, Ge S, Gu Y, He Z;  
 XX  
 DR WPI; 2002-427096/45.  
 XX  
 PT Polypeptide fragments of hepatitis E virus (HEV) for vaccine compositions  
 XX and diagnosis in preventing and treating HEV infection.  
 PS Example 1; Page 83-84; 91pp; Chinese.  
 XX  
 CC The sequence encodes hepatitis E virus (HEV) ORF-2 polypeptide. The  
 CC invention relates to novel n-mer polypeptides comprising amino acids of  
 CC ORF-2 of HEV where n = 2-180. The polypeptides of the invention have  
 CC virucide activity. The polypeptide of HEV, its fragments, and chimeric  
 CC proteins, are useful for vaccine compositions and diagnosis in preventing  
 CC and treating HEV infection by immunisation, which may also be applied in  
 CC the diagnosis of HEV infection  
 XX  
 SQ Sequence 1989 BP; 347 A; 666 C; 481 G; 495 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3.58e-103 Length: 1989  
 Score: 1095.00 Matches: 213  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0

```
US-10-089-292A-2 (1-213) x ABL60054 (1-1989)
Qy 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 20
Db 1180 CAGCTGTTCTACTCTCGTCCGCTGCTCTCAGCCAATGCGAGCGACTGTTAAGCTTTAT 1239
Qy 21 ThrSerValGluAsnAlaGlnAspLysGlyLeuAlaIleProHisAspIleAspLeu 40
Db 1240 ACATCTGTAGAGATGCTCAGCAGGATAAGGGTATTGCAATCCCGCATGACATCGACCTC 1299
Qy 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60
Db 1300 GGGAGTCTCGTGTAGTATTTCAGATTATGACCAACCAATGACAGACCGACCGACA 1359
Qy 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80
Db 1360 CTTTCCCGCAGCCCATCGGCCCTTTTCTGCTCCGAGCTAATGATGTCTTTGGCTT 1419
Qy 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100
Db 1420 TCTCTACCGCTGCGAGTATGACAGTCCACTTACGGCTCTTCGACCGCCGCTCTAT 1479
Qy 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120
Db 1480 GTCTCTGACTCTGTGACCTTGTAAATGTTGCGACCGCGCGAGCGCTTCCCGGTCA 1539
Qy 121 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnIleTyrSer 140
Db 1540 CTCGACTGGACCAAGTTCACACTTGATGTCGCCCTTTCCACCATCCAGCATTTTCA 1599
Qy 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160
Db 1600 AAGACCTTCTTTGTCTCTCGCGCTCCGCGTAAAGCTCTCTTTTGGAGGCGAGTACTACT 1659
Qy 161 LysAlaGlyTyrProTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180
Db 1660 AAGCCGGGTACCTTAAATTAATTAACCACTGCTAGTACCACTGCTGTTGAGAA 1719
Qy 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200
Db 1720 GCCGCTGGGCATCGGGTGTCTATTTCACCTTACACCATAGCTGGGTGGTCCCGTC 1779
Qy 201 SerLeuSerAlaValAlaValLeuAlaProProArg 213
Db 1780 TCTATTTCGGGGTGTGCTTTTGTAGCCCCCTCCGCGC 1818

RESULT 3
ADB97809
ID ADB97809 standard; DNA; 1989 BP.
XX AC ADB97809;
XX AC
XX DT 04-DEC-2003 (first entry)
XX DE HEV related DNA #SEQ ID 3.
XX KW Hepatitis E virus; HEV; virucide; vaccine; monoclonal antibody; NE2;
XX KW infection; ds.
XX OS Unidentified.
XX PN WO2003040187-A1.
XX PD 15-MAY-2003.
XX PF 08-NOV-2002; 2002WO-CN000797.
XX PR 08-NOV-2001; 2001CN-00134643.
XX PA (YANG-) YANG SHENG TANG CO LTD.
XX FI Xia N, Zhang J, Gu Y, Li S, Ge S, He Z;
```

XX WPI; 2003-441530/41.

Hepatitis E virus monoclonal antibodies and their active fragments, for use in diagnosis and developing remedies e.g. vaccines for preventing or treating hepatitis E virus infection.

Example 1; Page 133-135; 165pp; Chinese.

The invention relates to a monoclonal antibody that binds specifically to hepatitis E virus open-reading frame (ORF) 2. Also disclosed are nucleotide sequences encoding the heavy-chain and/or light-chain variable region of their degeneration sequences. The monoclonal antibody is selected from anti-HEV monoclonal antibody 8C11 secreted by hybridoma CTCC-C200116, anti-HEV monoclonal antibody 13D8 secreted by hybridoma CTCC-C200114, anti-HEV monoclonal antibody 8H3 secreted by hybridoma CTCC-C200117, and anti-HEV monoclonal antibody 16D7 secreted by hybridoma CTCC-C200114. The monoclonal antibodies and their active fragments are useful in diagnosis and developing remedies e.g. vaccines for preventing or treating hepatitis E virus infection. From the whole HEV gene obtained in patients, the HEV ORF fragment was prepared by using the method of Aye et al. The polypeptide NE2 was then expressed and isolated after biotechnological manipulations, which was characterised. Its analogs were also produced for immunising mice and construction of hybridomas. Tests were carried out to confirm usefulness of such polypeptides and monoclonal antibodies. The current sequence represents a HEV related DNA.

SQ Sequence 1989 BP; 347 A; 666 C; 481 G; 495 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	3.58e-103	Length:	1989
Score:	1095.00	Matches:	213
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-10-089-292A-2 (1-213) x ADB97809 (1-1989)

Qy 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 20

Db 1180 CAGCTGTTCTACTCTCGTCCGCTGCTCTCAGCCAATGCGAGCGACTGTTAAGCTTTAT 1239

Qy 21 ThrSerValGluAsnAlaGlnAspLysGlyLeuAlaIleProHisAspIleAspLeu 40

Db 1240 ACATCTGTAGAGATGCTCAGCAGGATAAGGGTATTGCAATCCCGCATGACATCGACCTC 1299

Qy 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60

Db 1300 GGGAGTCTCGTGTAGTATTTCAGATTATGACCAACCAATGACAGACCGACCGACA 1359

Qy 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80

Db 1360 CTTTCCCGCAGCCCATCGGCCCTTTTCTGCTCCGAGCTAATGATGTCTTTGGCTT 1419

Qy 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100

Db 1420 TCTCTACCGCTGCGAGTATGACAGTCCACTTACGGCTCTTCGACCGCCGCTCTAT 1479

Qy 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120

Db 1480 GTCTCTGACTCTGTGACCTTGTAAATGTTGCGACCGCGCGAGCGCTTCCCGGTCA 1539

Qy 121 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnIleTyrSer 140

Db 1540 CTCGACTGGACCAAGTTCACACTTGATGTCGCCCTTTCCACCATCCAGCATTTTCA 1599

Qy 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160

Db 1600 AAGACCTTCTTTGTCTCTCGCGCTCCGCGTAAAGCTCTCTTTTGGAGGCGAGTACTACT 1659

Qy 161 LysAlaGlyTyrProTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180

Db 1660 AAGCCGGGTACCTTAAATTAATTAACCACTGCTAGTACCACTGCTGTTGAGAA 1719

Qy 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200

Db 1720 GCCGCTGGGCATCGGGTGTCTATTTCACCTTACACCATAGCTGGGTGGTCCCGTC 1779

Qy 201 SerLeuSerAlaValAlaValLeuAlaProProArg 213

Db 1780 TCTATTTCGGGGTGTGCTTTTGTAGCCCCCTCCGCGC 1818



PR 05-APR-1990; 90US-00505888.  
 PR 05-APR-1991; 91US-00681078.  
 PR 17-JAN-1992; 92US-00822335.  
 PR 20-APR-1992; 92US-00870985.  
 PR 01-MAY-1992; 92US-00876941.  
 XX (GENE-) GENELABS TECHNOLOGIES INC.  
 XX  
 PI Tam AW, Reyes GR, Yarbrough PO;  
 XX  
 DR P-PSDB; AAW35820.  
 XX  
 PT Diagnosis of hepatitis E virus Burma and Mexico strain infection - by  
 PT immunoassay with hepatitis E virus peptide antigens.  
 XX  
 PS Disclosure; Col 29-32; 36pp; English.  
 XX  
 CC A method has been developed for detecting hepatitis E virus (HEV)  
 CC antibodies (Ab). The method comprises: (a) reacting a serum sample with a  
 CC HEV peptide antigen; and (b) examining the peptide for the presence of  
 CC bound Ab, where the presence of bound Ab indicates the presence of HEV  
 CC Ab. The present sequence encodes a specifically claimed HEV peptide  
 CC antigen. The method can be used to diagnose infection with the  
 CC enterically transmitted non-A/non-B viral hepatitis agent HEV,  
 CC specifically the HEV Burma and Mexico strains. (Updated on 25-MAR-2003 to  
 CC correct PF field.) (Updated on 17-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 984 BP; 180 A; 300 C; 246 G; 258 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1.3e-101 Length: 984  
 Score: 1076.00 Matches: 210  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.26% Indels: 0  
 DB: 2 Gaps: 0  
 US-10-089-292A-2 (1-213) x AAT96965 (1-984)  
 QY 1 GlnLeuPheTyrSerArgProValValSerAlaAenGlyGluProThrValLysLeuTyr 20  
 Db  
 181 CAGCTGTTCTACTCCGTCCTGGTGTCTCAGCAATGGCGAGCGACGCTTAAAGTTGAT 240  
 QY 21 ThrSerValGluAenAlaGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 40  
 Db 2A1 ACATCTGTAGAGATGCTCAGCAGGATAAGGTATTGCAATCCCGCATGACATTGACCTC 300  
 QY 41 GlyGluSerArgValValIleGlnAspTyrAspAenGlnHisGluGlnAspArgProThr 60  
 Db 301 GGAGAATCTCGTGTGGTATTTCAGGATTATGATAACCAACATGAACAAGATCGGCCGACG 360  
 QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80  
 Db 361 CTTTCTCCAGCCCCCATCGCCCTTCTCTGTCCTTCGAGCTAATGATGCTTTGGCTC 420  
 QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
 Db 421 TCTCTCACCGCTGCGGAGTATGACCACTTATGCTTCTTCGACTGGCCAGTTTAT 480  
 QY 101 ValSerAspSerValThrLeuValAenValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
 Db 481 GTTTCCTGACTCTGTGACCTTGGTTAATGTTGCGACCGCGCGCGGCGGTTGCCCGTGC 540  
 QY 121 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnLntYrSer 140  
 Db 541 CTCGATTGACCAAGGTACACTTCAGCGTCCGCCCTCTCTCCACCATCCAGCAGTACTCG 600  
 QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160  
 Db 601 AAGACTTCTTGTCTCGCGCTCCGCGTAAAGCTCTCTTCTTGGAGGAGGACACACT 660  
 QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAen 180

Db 661 AAAGCCGGGTACCTTATAATTATAACCACTGCTGACCACTGCTTGTCCAGAA 720  
 QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200  
 Db 721 GCCGCCGGCACCGGTGCTATTTCACCTTACCACTAGCCTGGGTGCTGGTCCCGTC 780  
 QY 201 SerIleSerAlaValAlaValLeuAlaPro 210  
 Db 781 TCCATTCTCGCGTTGCCGTTTTCAGCCCC 810  
 RESULT 6  
 AAT27109  
 ID AAT27109 standard; DNA: 1647 BP;  
 XX  
 AC AAT27109;  
 XX  
 DT 16-OCT-2003 (revised)  
 DT 06-AUG-1996 (first entry)  
 XX  
 DE Hepatitis E virus (Burma strain) 62K antigen DNA.  
 XX  
 KW HEV, enterically-transmitted non-A/non-B hepatitis virus; vaccine;  
 KW diagnosis; antigen; Spodoptera frugiperda; Sf9; insect; baculovirus;  
 KW capsid; db.  
 XX  
 OS Hepatitis E virus; Burma strain.  
 XX  
 PN WO9612807-A2.  
 XX  
 PD 02-MAY-1996.  
 XX  
 PF 23-OCT-1995; 95WO-US013703.  
 XX  
 PR 24-OCT-1994; 94US-00327952.  
 PR 13-OCT-1995; 95US-00542634.  
 XX  
 PA (GENE-) GENELABS TECHNOLOGIES INC.  
 XX  
 PI Fuerst TR, Mcatee CP, Yarbrough PO, Zhang Y;  
 XX  
 DR WPI; 1996-230608/23.  
 DR P-PSDB; AAR96090.  
 XX  
 PT Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as  
 PT diagnostic reagents for determining HEV infection and in vaccines.  
 XX  
 PS Disclosure; Page 72-73; 125pp; English.  
 XX  
 CC A DNA sequence (AAT27109) codes for 62K antigen (AAR96091) comprising the  
 CC C-terminal 549 amino acids of hepatitis E virus (HEV) Burma strain capsid  
 CC protein (AAR96089). It is obtd. by PCR amplification of HEV Burma ORF-2  
 CC (AAT27107) from pBEr1 plasmid. The amplified fragment is inserted into  
 CC vector pGSX for prodn. 62K in E. coli cells. Expression of full-length  
 CC ORF-2 in Sf9 insect cells using a baculovirus system also results in  
 CC prodn. of 62K antigen (see also AAR96101 and AAR96103). 62K represents an  
 CC improved antigen, in comparison to bacterial expressed proteins, for use  
 CC in HEV diagnostic assays, and also has excellent immunogenic properties.  
 CC (Updated on 16-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 1647 BP; 302 A; 533 C; 391 G; 421 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2.58e-101 Length: 1647  
 Score: 1076.00 Matches: 210  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 98.26% Indels: 0  
 DB: 2 Gaps: 0  
 US-10-089-292A-2 (1-213) x AAT27109 (1-1647)  
 QY 1 GlnLeuPheTyrSerArgProValValSerAlaAenGlyGluProThrValLysLeuTyr 20

Db 847 CAGCTGTTCTACTCCGCTCCGCTGTTCTCAGCAATGGCGAGCGACTGTTAAGTTGAT 906  
QY 21 ThrSerValGluAunAlaGlnAspLysGlyLeAlaIleProHisAspIleAspLeu 40  
Db 907 ACATCTGTAGAGATGCTCAGCAGGATAAGGATATTGCAATCCCGCATGACATTGACCTC 966  
QY 41 GlyGluSerArgValAlaIleGlnAspTyrAspAenGlnHisGluGlnAspArgProThr 60  
Db 967 GGAGAATCTCGTGTGTTATTGAGATTATGATAACCAACATGAACAAGATCGGCCGACG 1026  
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80  
Db 1027 CTTCTCCAGCCCATCGCCCTTCTCTGCTCCAGCTAATGATGCTTTGGCTC 1086  
QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrThrGlySerSerThrGlyProValTyr 100  
Db 1087 TCTCTCAGCGCTGCGGAGTATGACGATGACCTTATGCTCTTCGACTGGCCAGTTTAT 1146  
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
Db 1147 GTTCTGACTCTGTGACCTTGTGTTAATGTTGCGACCGCGCGGCTTTCGCCGGTGC 1206  
QY 121 LeuAspTyrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140  
Db 1207 CTCGATTGACCAAGTGCACCTTGACGGTGCCTCCCTCTCCACCATCCAGCAGTACTCG 1266  
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160  
Db 1267 AAGACCTTCTTGTCTGCTCCGCTCCGCGGTAAGCTCTCTTCTGGGAGCGACCAACT 1326  
QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180  
Db 1327 AAAGCCGGGTACCTTATTAATTAACACCATGCTAGCGACCAACTGCTTGTGCGAAT 1386  
QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200  
Db 1387 GCGCGCGGCGACCGGGTGCCTATTTCCACTTACACCATAGCTGGTGGTCTGGTCCCGTC 1446

## RESULT 7

AAV61686  
ID AAV61686 standard; DNA; 1647 BP.  
XX  
AC AAV61686;  
XX  
DT 03-DEC-1998 (first entry)  
DE Hepatitis E virus hollow particle protein ORF2 DNA.  
XX  
KW Hollow particle protein; virus; antibody; detection; immunoassay;  
KW infection; ss.  
XX  
OS Hepatitis virus.  
XX  
FH Key Location/Qualifiers  
FT CDS 1..1647  
FT FT /\*tag= a  
FT FT /product= "hollow particle protein"  
FT FT /note= "No start codon given for ORF2"  
XX  
FN JF10234393-A.  
XX  
PD 08-SEP-1998.  
XX  
PF 28-FEB-1997; 97JP-00062445.  
XX  
PR 28-FEB-1997; 97JP-00062445.  
XX  
PA (DENK-) DENKA SEIKEN KK.

PA (KOKU-) KOKURITSU YOBO EISEI KENKYUSHO.  
XX  
DR WPI: 1998-535037/46.  
DR P-PSDB; AAW76367.  
XX  
PT Hepatitis E virus hollow particle poly peptide(s) and nucleic acids  
PT encoding it - useful for more accurate detection of HEV in samples, using  
PT immuno-assays and nucleic acid hybridisation.  
XX  
PS Claim 1; Page 15-17; 29pp; Japanese.  
XX  
CC This sequence encodes a Hepatitis E viral hollow particle protein  
CC described as ORF2. This polypeptide can be used to raise antibodies to  
CC detect HEV infection in samples, e.g. by immuno-assay based techniques,  
CC and the nucleic acid can be used for the same in nucleic acid  
CC hybridisation assays. The polypeptide and nucleic acid allow more  
CC accurate detection of HEV than previously possible  
XX  
SQ Sequence 1647 BP; 303 A; 535 C; 390 G; 419 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2,58e-101 Length: 1647  
Score: 1076.00 Matches: 210  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.26% Indels: 0  
DB: 2 Gaps: 0

US-10-089-292A-2 (1-213) X AAV61686 (1-1647)

QY 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 20  
Db 847 CAGCTGTTCTACTCCGCTCCGCTGTTCTCGGCCAATGGCGAGCGACTGTTAAGTTGAT 906  
QY 21 ThrSerValGluAunAlaGlnAspLysGlyLeAlaIleProHisAspIleAspLeu 40  
Db 907 ACATCTGTAGAGATGCTCAGCAGGATAAGGATATTGCAATCCCGCATGACATTGACCTC 966  
QY 41 GlyGluSerArgValValIleGlnAspTyrAspAenGlnHisGluGlnAspArgProThr 60  
Db 967 GGAGAATCTCGTGTGTTATTGAGATTATGACACCAACATGAACAAGATCGGCCGACG 1026  
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80  
Db 1027 CTTCTCCAGCCCATCGGCCCTTCTCTGCTTCGAGCTAATGATGCTTTGGCTC 1086  
QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
Db 1087 TCTCTCAGCGCTGCGGAGTATGACCATTCATGCTCTTCGACTGGCCAGTTTAT 1146  
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
Db 1147 GTTCTGACTCTGTGACCTTGTGTTAATGTTGCGACCGCGCGCGCTTTCGCCGGTGC 1206  
QY 121 LeuAspTyrThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140  
Db 1207 CTCGATTGACCAAGTGCACCTTGACGGTGCCTCCCTCTCCACCATCCAGCAGTACTCG 1266  
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160  
Db 1267 AAGACCTTCTTGTCTGCTCCGCTCCGCGGTAAGCTCTCTTCTGGGAGCGACCAACT 1326  
QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180  
Db 1327 AAAGCCGGGTACCTTATTAATTAACACCATGCTAGCGACCAACTGCTTGTGCGAAT 1386  
QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200  
Db 1387 GCGCGCGGCGACCGGGTGCCTATTTCCACTTACACCATAGCTGGTGGTCTGGTCCCGTC 1446  
QY 201 SerIleSerAlaValAlaValLeuAlaPro 210  
Db 1447 TCCATTTCTGCGGTGCTGTTTGTAGCCCCC 1476



```

RESULT 8
AAAT96959
ID AAT96959 standard; DNA; 1983 BP.
XX
AC AAT96959;
XX
DT 17-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 26-FEB-1998 (first entry)
XX
DE Hepatitis E virus Burma strain ORF2.
XX
KW Hepatitis E virus; Burma; Mexico; immunoassay; peptide antigen; antibody;
KW diagnosis; HEV; ds.
XX
OS Hepatitis E virus; - Burma strain.
XX
FH Key Location/Qualifiers
FT CDS 1..1983
FT FT /*tag= a
FT FT /product= "ORF2_protein"
XX
PN US5686239-A.
XX
PD 11-NOV-1997.
XX
PF 09-MAY-1994; 94US-00240049.
XX
PR 17-JUN-1988; 88US-00208997.
PR 11-APR-1989; 89US-00336672.
PR 16-JUN-1989; 89US-00367486.
PR 13-OCT-1989; 89US-00420921.
PR 05-APR-1990; 90US-00505888.
PR 05-APR-1991; 91US-00681078.
PR 17-JAN-1992; 92US-00822335.
PR 20-APR-1992; 92US-00870985.
PR 01-MAY-1992; 92US-00876941.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Tam AW, Reyes GR, Yarbough PO;
XX
DR WPI; 1997-558132/51.
DR P-PSDB; AAW35826.
XX
PT Diagnosis of hepatitis E virus Burma and Mexico strain infection - by
PT immunoassay with hepatitis E virus peptide antigens.
XX
PS Disclosure; Col 23-24; 36pp; English.
XX
CC A method has been developed for detecting hepatitis E virus (HEV)
CC antibodies (Ab). The method comprises: (a) reacting a serum sample with a
CC HEV peptide antigen; and (b) examining the peptide for the presence of
CC bound Ab, where the presence of bound Ab indicates the presence of HEV
CC Ab. The present sequence encodes the protein from the open reading frame,
CC ORF2, from HEV Burma strain. The method can be used to diagnose infection
CC with the enterically transmitted non-A/non-B viral hepatitis agent HEV,
CC specifically the HEV Burma and Mexico strains. (Updated on 25-MAR-2003 to
CC correct PF field.) (Updated on 17-OCT-2003 to standardise OS field)
XX
SQ Sequence 1983 BP; 340 A; 665 C; 486 G; 492 T; 0 U; 0 Other;

Alignment Scores:
Pred No.: 3 29e-101 Length: 1983
Score: 1076.00 Matches: 210
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.26% Indels: 0
DB: 2 Gaps: 0

US-10-089-292A-2 (1-213) x AAT96959 (1-1983)

```

---

```

Qy 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValIysLeuTyr 20
Db 1180 CAGCTGTTCTACTCCCGTCCCGTTGTCTCAGCCAATGCGAGCCGACTGTTAAGTTGAT 1239
Qy 21 ThrSerValGluAsnAlaGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 40
Db 1240 ACATCTGTAGAGAAATGCTCAGCAGGATAGGGTATTGCAATCCCGCATGACATTGACCTC 1299
Qy 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60
Db 1300 GGAGAATCTCGTGTGGTTATTTCAGGATTATGATACCAACAATGAACAGATCGGCCGACG 1359
Qy 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80
Db 1360 CTTTCTCAGCCCCCATCGGCCCTTTCTCTGCTCTCGAGCTAATGATGTGTTGGCTC 1419
Qy 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100
Db 1420 TCTCTCACCCTGCGAGATATGACCACTTATGGCTCTTCGACTGGCCAGTTTAT 1479
Qy 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120
Db 1480 GTTTCGACTCTGTGACCTTGGTTAATGTTGCGACCGCGCGCAGGCGGTTGCCCGGTG 1539
Qy 121 LeuAspTrpThrIysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140
Db 1540 CTCGATTGGACCAAGGTCACACTTTCGCGTCCGCCCTCTCCACATCCAGCAGTACTCG 1599
Qy 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTyrGluAlaGlyThrThr 160
Db 1600 AAGACCTTCTTGTCTCGCGCTCCGCGGTAAAGTCTCTTCTGGAGGAGGAGGACAACT 1659
Qy 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuLeuValGluAsn 180
Db 1660 AAAGCCGGGTACCTTATTAATTAAACACCACTGCTAGCAGCAACTGCTTGTGCGAAT 1719
Qy 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProVal 200
Db 1720 GCCCGCGGACCGGGTGCCTATTTCACATTACACCACTAGCCTGGGTGGTGGTCCCGTC 1779
Qy 201 SerIleSerAlaValAlaValLeuAlaPro 210
Db 1780 TCCATTCTCGCGTTGCGGTTTATAGCCCCC 1809
XX
RESULT 9
AAV61688
ID AAV61688 standard; DNA; 1983 BP.
XX
AC AAV61688;
XX
DT 03-DEC-1998 (first entry)
DE Hepatitis E virus hollow particle protein DNA #2.
XX
KW Hollow particle protein; virus; antibody; detection; immunoassay;
XX infection; ss.
XX
OS Hepatitis virus.
XX
PN JP10234383-A.
XX
PD 08-SEP-1998.
XX
PF 28-FEB-1997; 97JP-00062445.
XX
PR 28-FEB-1997; 97JP-00062445.
XX
PA (DENK-) DENKA SEIKEN KK.
XX
PA (KOKU-) KOKURITSU YOBO EISEI KENKYUSHO.
XX
DR WPI; 1998-535037/46.
DR P-PSDB; AAW76369.
XX

```

PT Hepatitis E virus hollow particle poly:peptide(s) and nucleic acids  
PT encoding it - useful for more accurate detection of HEV in samples, using  
PT immuno-assays and nucleic acid hybridisation.

```
Db 1420 TCTCTACCGCTGGCGAGTATGACAGCTCCACTTACGGCTCTTCGACCGGCCAGTCTAT 1479
Qy 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120
Db 1480 GTCTCTGACTCTGTGACCTTGGTTAATGTTGGACCGCGCGAGCGCTTGGCCGGTCA 1539
Qy 121 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140
Db 1540 CTCGACTGGACCAAGGTACACACTTGTATGTCGCCCTTTCCACCATCCAGCAGTATTCA 1599
Qy 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160
Db 1600 AAGACCTTCTTTGTCTCCGCTCCGCGTAAAGTCTCTCTTTGGAGCGAGGACTACT 1659
Qy 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180
Db 1660 AAGCGCGGTACCTTATATTAACACCACTGCTAGTACCACTGCTCTGTTGAGAA 1719
Qy 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200
Db 1720 GCCGCTGGGCATCGGGTTCCTATTTCACCTTACACCACTAGCCTGGGTGCTGCCGTC 1779
Qy 201 SerIleSerAlaValAlaValLeuAlaPro 210
Db 1780 TCTATTTCGCGGTGCTGTTTAGCCCCC 1809

RESULT 11
ABL60053
ID ABL60053 standard; DNA; 1990 BP.
XX
AC ABL60053;
XX
XX 16-AUG-2002 (first entry)
XX
DE Hepatitis E virus ORF-2 DNA #1.
XX
XX Hepatitis E virus; ORF-2; HEV; virucide; immunisation; gene; ds.
XX
XX Hepatitis E virus.
XX
XX WO200240681-AL.
XX
PD 23-MAY-2002.
XX
PF 30-SEP-2001; 2001WO-CN001459.
XX
PR 30-SEP-2000; 2000CN-00130634.
XX
FA (YANG-) YANG SHENG TANG CO LTD.
XX
PI Xia N, Zhang J, Li S, Ge S, Gu Y, He Z;
XX
XX WPI; 2002-427096/45.
DR
XX
PT Polypeptide fragments of hepatitis E virus (HEV) for vaccine compositions
XX and diagnosis in preventing and treating HEV infection.
XX
XX Example 1; Page 82-83; 91pb; Chinese.
XX
CC The sequence encodes hepatitis E virus (HEV) ORF-2 polypeptide. The
XX invention relates to novel n-mer polypeptides comprising amino acids of
XX ORF-2 of HEV where n = 2-180. The polypeptides of the invention have
XX virucide activity. The polypeptide of HEV, its fragments, and chimeric
XX proteins, are useful for vaccine compositions and diagnosis in preventing
XX and treating HEV infection by immunisation, which may also be applied in
XX the diagnosis of HEV infection
XX
SQ Sequence 1990 BP; 348 A; 666 C; 481 G; 495 T; 0 U; 0 Other;
```

Alignment Scores:

Pred. No.: 3.31e-101 Length: 1990

Score: 1076.00 Matches: 210

```
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.26% Indels: 0
DB: 6 Gaps: 0

US-10-089-292A-2 (1-213) x ABL60053 (1-1990)
Qy 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 20
Db 1180 CAGCTGTTTACTCTCGTCCCGTCTCTCAGCCAATGGCGAGCGACTGTTAAGCTTTAT 1239
Qy 21 ThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 40
Db 1240 ACATCTGTAGAGATGCTCAGCAGATTAAGGGTATTGCAATCCCGCATGACATCCACCTC 1299
Qy 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60
Db 1300 GGGGAGTCTCGTGTAGTTATTTCAGAGTATTATGACAACCAACATGAGCAGACCGACGACA 1359
Qy 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80
Db 1360 CTTTCCCGCAGCCCATCGCGCCCTTTTCTGTCTCCGAGCTAATGATGTGCTTTGGCTT 1419
Qy 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100
Db 1420 TCTCTCACCGCTGCCGAGTATGACCACTTACGGCTCTTCGACCGGCCAGTCTAT 1479
Qy 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120
Db 1480 GTCTCTGACTCTGTGACCTTGGTTAATGTTGGACCGCGCGAGCGCTTGGCCGGTCA 1539
Qy 121 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140
Db 1540 CTCGACTGGACCAAGGTACACACTTGTATGTCGCCCTTTCCACCATCCAGCAGTATTCA 1599
Qy 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160
Db 1600 AAGACCTTCTTTGTCTCCGCTCCGCGTAAAGTCTCTCTTTGGAGCGAGGACTACT 1659
Qy 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180
Db 1660 AAGCGCGGTACCTTATATTAACACCACTGCTAGTACCACTGCTCTGTTGAGAA 1719
Qy 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200
Db 1720 GCCGCTGGGCATCGGGTTCCTATTTCACCTTACACCACTAGCCTGGGTGCTGCCGTC 1779
Qy 201 SerIleSerAlaValAlaValLeuAlaPro 210
Db 1780 TCTATTTCGCGGTGCTGTTTAGCCCCC 1809

RESULT 12
ADB97808
ID ADB97808 standard; DNA; 1990 BP.
XX
AC ADB97808;
XX
XX 04-DEC-2003 (first entry)
XX
DE HEV ORF2 DNA #SEQ ID 2.
XX
XX Hepatitis E virus; HEV; virucide; vaccine; monoclonal antibody; NE2;
XX infection; ds.
XX
XX Hepatitis E virus.
XX
XX WO2003040187-AL.
XX
PD 15-MAY-2003.
XX
PF 08-NOV-2002; 2002WO-CN000797.
XX
XX 08-NOV-2001; 2001CN-00134643.
```

XX (YANG-) YANG SHENG TANG CO LTD.  
XX Xia N, Zhang J, Gu Y, Li S, Ge S, He Z;  
XX WPI; 2003-441530/41.  
XX Hepatitis E virus monoclonal antibodies and their active fragments, for  
PT use in diagnosis and developing remedies e.g. vaccines for preventing or  
PT treating hepatitis E virus infection.  
XX Example 1; Page 131-133; 165pp; Chinese.  
XX The invention relates to a monoclonal antibody that binds specifically to  
XX hepatitis E virus open-reading frame (ORF) 2. Also disclosed are  
XX nucleotide sequences encoding the heavy-chain and/or light-chain variable  
XX region of their degeneration sequences. The monoclonal antibody is  
XX selected from anti-HEV monoclonal antibody 8C11 secreted by hybridoma  
XX C2TC-C200116, anti-HEV monoclonal antibody 13D8 secreted by hybridoma  
XX C2TC-C200114, anti-HEV monoclonal antibody 8H3 secreted by hybridoma  
XX C2TC-C200117, and anti-HEV monoclonal antibody 16D7 secreted by  
XX hybridoma C2TC-C200114. The monoclonal antibodies and their active  
XX fragments are useful in diagnosis and developing remedies e.g. vaccines  
XX for preventing or treating hepatitis E virus infection. From the whole  
XX HEV gene obtained in patients, the HEV ORF fragment was prepared by using  
XX the method of Aye et al. The polypeptide NE2 was then expressed and  
XX isolated after biotechnological manipulations, which was characterised.  
XX Its analogs were also produced for immunising mice and construction of  
XX hybridomas. Tests were carried out to confirm usefulness of such  
XX polypeptides and monoclonal antibodies. The current sequence represents  
XX the HEV ORF2 encoding DNA.  
XX Sequence 1990 BP; 348 A; 666 C; 481 G; 495 T; 0 U; 0 Other;  
SQ

Alignment Scores:  
Pred. No.: 3,31e-101 Length: 1990  
Score: 1076.00 Matches: 210  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.26% Indels: 0  
DB: 10 Gaps: 0

US-10-089-292A-2 (1-213) x ADB97808 (1-1990)

QY 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyLupProThrValLysLeuTyr 20  
DB 1180 CAGCTGTTTACTCTCTGTCCTCCGTCGTCAGCAATGGCGAGCCGACTGTTAAGCTTTAT 1239

QY 21 ThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspLeu 40  
DB 1240 ACATCTGTAGAGAATGCTCAGCAGGATAAGGGTATTGCAATCCGCGATGACATCGACCTC 1299

QY 41 GlyLysSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60  
DB 1300 GGGGACTCTCGTGTAGTTATTTCAGGAATTATGCAACCAATGAGCAGACCGACCGACA 1359

QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80  
DB 1360 CCTTCCCGAGCCCATCGGCGCCCTTTCTGTCCTCGAGCTAATGATGTGCTTGGCTT 1419

QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
DB 1420 TCTCTCACCGCTCCGAGTATGACCATGTCACCTACGGCTCTTCGACCGCGCCGCTAT 1479

QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120  
DB 1480 GTCTCTGACTCTGTGACCTTGGTTATGTTGCCAGCGCGCGAGCCGTTGCCCGGTCA 1539

QY 121 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnIntyrSer 140  
DB 1540 CTCGACTGGACCAAGGTCACATGATGGTCGCCCCCTTTCCACCATCCAGCAGTATTCA 1599

QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160

DB 1600 AAGACCTTCTTTGTCTCTCCGCGTAAAGCTCTCTTTTGGGAGGCGAGGTACTACT 1659

QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180

DB 1660 AAGCCGGGTACCCCTTATAATTTAAACACCACTGCTAGTACCACTGCTGTTGAGAT 1719

QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrSerLeuGlyAlaGlyProVal 200

DB 1720 GCGCTGGGCATCGGTTGCTATTTCACATTACCACTAGCTGGGTGGTGGTCCCGTC 1779

QY 201 SerIleSerAlaValAlaValLeuAlaPro 210

DB 1780 TCTATTTCGCGGTGCTGCTTTTAGCCCCC 1809

RESULT 13  
AAT27107  
ID AAT27107 standard; DNA; 2049 BP.  
XX  
AC AAT27107;  
XX  
DT 16-OCT-2003 (revised)  
DT 06-AUG-1996 (first entry)  
XX  
DE Hepatitis E virus (Burma strain) ORF-2.  
XX  
KW HEV; enterically-transmitted non-A/non-B hepatitis virus; vaccine;  
KW diagnosis; antigen; Spodoptera frugiperda; Sf9; insect; baculovirus;  
KW capsid; ds.  
XX  
OS Hepatitis E virus; Burma strain.  
XX  
PN WO9612807-A2.  
XX  
PD 02-MAY-1996.  
XX  
PF 23-OCT-1995; 95WO-US013703.  
XX  
PR 24-OCT-1994; 94US-00327952.  
PR 13-OCT-1995; 95US-00542634.  
XX  
PA (GENE-) GENELABS TECHNOLOGIES INC.  
XX  
PI Fuerst TR, Mcatee CP, Yarbrough PO, Zhang Y;  
XX  
DR WPI; 1996-230608/23.  
XX  
PS P-PSDB; AAR96089.  
XX  
PT Hepatitis E virus (HEV) antigens derived from ORF 2 - useful as  
PT diagnostic reagents for determining HEV infection and in vaccines.  
XX  
PS Disclosure; Page 69-70; 125pp; English.  
XX  
CC The hepatitis E virus (HEV) Burma strain ORF-2 (AAT27107) codes for the  
CC putative capsid protein (AAR96089). PCR amplification of ORF-2 allows  
CC prodn. of capsid protein or of C-terminal regions of the capsid protein  
CC (see also AAR96091, AAR96093 and AAR96095) and expression in Spodoptera  
CC frugiperda Sf9 insect cells using a baculovirus vector provides  
CC recombinant C-terminal regions (see also AAR96101 and AAR96103) useful as  
CC diagnostic reagents and in vaccines. The HEV Mexico strain ORF-2  
CC (AAT27108) may similarly be used. (Updated on 16-OCT-2003 to standardise  
CC OS field)  
XX  
SQ Sequence 2049 BP; 344 A; 686 C; 496 G; 523 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3,44e-101 Length: 2049  
Score: 1076.00 Matches: 210  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.26% Indels: 0  
DB: 2 Gaps: 0

```
US-10-089-292A-2 (1-213) x ANT27107 (1-2049)
QY 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValIysLeuTyr 20
Db 1180 CAGCTGTTCTACTCCGCTCCGCTGCTCTCAGCAATGGCGAGCCGACTGTTAAGTTGAT 1239
QY 21 ThrSerValGluAsnAlaGlnAspIysGlyIleAlaIleProHisAspIleAspLeu 40
Db 1240 ACATCTGTAGAGAAATGCTCAGCAGGATAAGGGATTGCAATCCCGCATGCACATTGACCTC 1299
QY 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60
Db 1300 GGAGAAATCTCGTGTGTTATTACAGATTATGATATACCAACATGACAAAGATCGGCCGAG 1359
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80
Db 1360 CCTTCTCCAGCCCATCGCCCTTCTCTGTCCTTCGAGCTAATGATGTGCTTTGGCTC 1419
QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100
Db 1420 TCTCTCACCGCTGCCGAGTATGACCACTTATGCTCTTCGACTGGCCAGTTTAT 1479
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyValAlaGlnAlaValAlaArgSer 120
Db 1480 GTTCTGACTCTGTGACCTTGGTTAATGTTGCACCGCGCGCAGCGCTGTCGCCGTCG 1539
QY 121 LeuAspTTPThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnIleNTrpSer 140
Db 1540 CTCGATTGGACCAAGGTCACACTTCAGCGTCGCCCTCTCCACCATCCAGCAGTACTCG 1599
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160
Db 1600 AAGACCTTCTTGTCTCTCGCGTCCGCGGTAAGCTCTCTTCTGGGAGCGAGGCACAACT 1659
QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180
Db 1660 AAAGCGGGTACCTTAAATTAATTAACCACTGTAGGACCAACTGCTGTGTCGAGAA 1719
QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200
Db 1720 GCCGCGGGCAGCGGTCGCTATTTCACCTTACACCACTAGGCTGGGTGGTCCCGTC 1779
QY 201 SerIleSerAlaValAlaValLeuAlaPro 210
Db 1780 TCCATTTCGCGGTGCGGTTTATAGCCCCC 1809
RESULT 14
AAS01448
ID AAS01448 standard; cDNA; 2054 BP.
AC AAS01448;
XX
XX
XX 11-SEP-2003 (revised)
XX 18-JUL-2001 (first entry)
DE Hepatitis E virus (HEV) ORF2 cDNA.
XX
XX Hepatitis E virus; HEV; E2; pE2; ORF2; immunoreactive viral peptide;
KW antigenic activity; viral infection; ss.
XX
XX Hepatitis E virus; Chinese strain D11092.
OS
FH Key Location/Qualifiers
FT CDS 1..1983
FT FT /*tag= a
FT FT /product= "Major HEV structural protein"
FT FT 1180..1190
FT FT /*tag= b
FT FT /note= "Corresponds to residues 1-810 of E2"
FT FT replace(1810..1812, CC)
FT FT /*tag= c
FT FT /note= "Single base deletion present in E2"
XX
```

FN WO200122916-A2.  
XX  
PD 05-APR-2001.  
XX  
PF 28-SEP-2000; 2000WO-IB0011393.  
XX  
PR 30-SEP-1999; 99CA-02283538.  
XX  
PA (UYHK-) UNIV HONG KONG.  
XX  
PI Ng MH, Im S, Zhang J;  
XX WPI; 2001-266109/27.  
XX  
PT New immunoreactive peptide pE2, cloned from the genome of a Chinese  
PT strain of hepatitis E virus (HEV), for detecting and diagnosing HEV  
PT infection and for use as a vaccine for prevention of the viral infection  
PT in humans.  
XX  
PS Disclosure; Fig 2A-2D; 101pp; English.  
XX  
CC The present sequence for hepatitis E virus (HEV) ORF2 cDNA is used to  
CC clone a cDNA fragment E2 (AAS01434) which encodes for a novel peptide,  
CC pE2. During the cloning of the E2 fragment a single base deletion  
CC (presumed to be a PCR amplification error) at position 1811 of ORF2  
CC resulted in a frame shift causing the premature termination at a new stop  
CC codon at position 1820-1822, and giving rise to a smaller peptide, pE2.  
CC Peptide pE2 is a highly immunoreactive viral peptide which can be used in  
CC the development of diagnostic methods for the detection of HEV. Peptide  
CC pE2 which has antigenic activity is also useful as a vaccine for  
CC immunising an individual against infection from HEV and for determining  
CC the presence of HEV antibodies in a biological test sample such as human  
CC blood serum or plasma for detection and diagnosis of HEV infection. An  
CC antibody against pE2 is useful for detecting an HEV particle, in a  
CC biological test sample. The use of recombinant peptides such as pE2  
CC peptide have advantage over the prior art of using live intact virus  
CC particles by avoiding the potential risk of infectivity. The ORF3  
CC (AAS01449) of HEV Chinese strain D11092 is used to clone a cDNA fragment  
CC E3 (AAS001450) which encodes pE3 peptide. (Updated on 11-SEP-2003 to  
CC standardise OS field)  
XX  
SQ Sequence 2054 BP; 355 A; 687 C; 491 G; 521 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3.45e-101 Length: 2054  
Score: 1076.00 Matches: 210  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 98.26% Indels: 0  
DB: 4 Gaps: 0

US-10-089-292A-2 (1-213) x AAS01448 (1-2054)  
QY 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValIysLeuTyr 20  
Db 1180 CAGCTGTTCTACTCCGCTCCGCTGCTCTCAGCAATGGCGAGCCGACTGTTAAGTTTAT 1239  
QY 21 ThrSerValGluAsnAlaGlnAspIysGlyIleAlaIleProHisAspIleAspLeu 40  
Db 1240 ACATCTGTAGAGAAATGCTCAGCAGGATAAGGGATTGCAATCCCGCATGCACCTC 1299  
QY 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60  
Db 1300 GGCGAGTCTCGTGTAGTTATTTCAGGATTATGACAAACCAACATGACGAGCCGACGACA 1359  
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80  
Db 1360 CTTTCCCGAGCCCATCGGCCCTTTTCTGCTCTCCGAGCTAATGATGTGCTTTGGCTT 1419  
QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100  
Db 1420 TCTCTCACCGCTGCCGAGTATGACCACTTATGCTCTTCGACTGGCCAGTTTAT 1479

```
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120
Db 1480 GTCTCTGACTCTGTGACCTTGGTTAATGTTGCGACCGCGCGAGCGGTTGCCCGGTCA 1539
QY 121 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140
Db 1540 CTCGACTGACCAAGGTACACTTGTATGTCGCCCTTTCCACCATCCAGCAGTATTCA 1599
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160
Db 1600 AAGACTCTTCTTGTCTCGCGCTCCGCGTAGACTCTCTTTTGGAGGAGGACTACTACT 1659
QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180
Db 1660 AAGCGCGGTACCTTATAATTATACACCACTGCTAGTGACCAACTGCTGTTGAGAAT 1719
QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200
Db 1720 GCCGCTGGGCATCGGGTTGCTATTTCCTTACACCACTAGCGTGGGTGCTGGTCCCGTC 1779
QY 201 SerIleSerAlaValAlaValLeuAlaPro 210
Db 1780 TCTATTCCGCGGTGCTGTTTTAGCCCCC 1809
RESULT 15
ID AAQ47129 standard; DNA; 2094 BP.
XX AAQ47129;
AC AAQ47129;
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 11-JAN-1994 (first entry)
DE HEV third and second ORFs.
XX Enterically transmitted non-A non-B hepatitis; ET-NANB; vaccine; ss.
XX Hepatitis E virus; Burma strain.
XX Key Location/Qualifiers
FT CDS 5..376
FT /tag= a
FT /label= ORF3
FT 46..2094
FT /tag= b
FT /label= ORF2
FT 275..376
FT /tag= c
FT /label= 406.4-2
FT 718..2029
FT /tag= d
FT /label= C2
FT 1045..2029
FT /tag= e
FT /label= SG3
FT 1882..2029
FT /tag= f
FT /label= 406.3-2
XX WO9314116-AL.
XX 22-JUL-1993.
XX 15-JAN-1993; 93WO-US000459.
XX 17-JAN-1992; 92US-0082335.
XX 01-MAY-1992; 92US-00876941.
XX (GENE-) GENELABS TECHNOLOGIES INC.
XX (USSH ) US SEC DEPT HEALTH.
XX Reyes GR, Bradley DW, Tam AW, Carl M;
```

```
XX WPI; 1993-243144/30.
DR P-PSDB; AAR38784, AAR38785.
XX New immunogenic hepatitis E virus (HEV) peptide(s) - are from the ORF1,
PT ORF2 and ORF3 regions of HEV, useful as a vaccine against HEV infection.
XX Disclosure; Fig 7; 48pp; English.
XX Immunogenic hepatitis E virus (HEV) peptides are selected from the ORF1,
CC ORF2 and ORF3 regions of HEV. The peptides can be used in vaccines to
CC prevent infection by HEV. The antibodies can neutralise and block HEV
CC infection and can be used to prevent or treat HEV infection. The peptides
CC and antibodies can also be used as diagnostic reagents. (Updated on 25-
CC MAR-2003 to correct PN field.) (Updated on 24-OCT-2003 to standardise OS
CC field)
XX Sequence 2094 BP; 353 A; 697 C; 508 G; 536 T; 0 U; 0 Other;
SQ Alignment Scores:
Pred. No.: 3.54e-101 Length: 2094
Score: 1076.00 Matches: 210
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 98.26% Indels: 0
DB: 2 Gaps: 0
US-10-089-292A-2 (1-213) x AAQ47129 (1-2094)
QY 1 GlnLeuPheTyrSerArgProValValSerAlaAsnGlyGluProThrValLysLeuTyr 20
Db 1225 CAGCTGTTCTACTCCCGTCCGTTGTCTCAGCAATGGCGAGCCGACTGTTAAGTTGAT 1284
QY 21 ThrSerValGluAsnAlaGlnGlnAspLysGlyIleAlaIleProHisAspIleAspLeu 40
Db 1285 ACATCTGTAGAGATGCTCAGCAGGATAAGGTATTGCAATCCCGCATGACATTGACCTC 1344
QY 41 GlyGluSerArgValValIleGlnAspTyrAspAsnGlnHisGluGlnAspArgProThr 60
Db 1345 GGAGAAATCTCGTGTGTTTATTTCAGGATTATGATATACCAACATGAACAAGATCGGCCGAG 1404
QY 61 ProSerProAlaProSerArgProPheSerValLeuArgAlaAsnAspValLeuTrpLeu 80
Db 1405 CCTTCTCCAGCCCCATCGGCCCTTCTCTGCTTCGAGCTAATGATGTTGGCTC 1464
QY 81 SerLeuThrAlaAlaGluTyrAspGlnSerThrTyrGlySerSerThrGlyProValTyr 100
Db 1465 TCTCTCACGCTGCCGAGTATGACCACTTATGCTCTTCGACTGGCCAGTTTAT 1524
QY 101 ValSerAspSerValThrLeuValAsnValAlaThrGlyAlaGlnAlaValAlaArgSer 120
Db 1525 GTTCTGACTCTGTGACCTTGGTTAATGTTGCGACCGCGCGCGGCTGTCGCCGCTG 1584
QY 121 LeuAspTrpThrLysValThrLeuAspGlyArgProLeuSerThrIleGlnGlnTyrSer 140
Db 1585 CTCGATTGGACCAAGGTACACTTGACCGTCCGCCCTCTCCACCATCCAGCAGTACTCG 1644
QY 141 LysThrPhePheValLeuProLeuArgGlyLysLeuSerPheTrpGluAlaGlyThrThr 160
Db 1645 AAGACTTCTTGTCTCTCGCGCTCCGCGTAAAGTCTCTTTCTGGGAGGAGGCAAACT 1704
QY 161 LysAlaGlyTyrProTyrAsnTyrAsnThrThrAlaSerAspGlnLeuValGluAsn 180
Db 1705 AAAGCGGGTACCTTATAATTATACACCACTGTAGCAGCAACTGCTGTCAGAAAT 1764
QY 181 AlaAlaGlyHisArgValAlaIleSerThrTyrThrThrSerLeuGlyAlaGlyProVal 200
Db 1765 GCCGCGGGCACCGGGTGGCTATTATTTCCACTTACACCACTAGCTGGGTGCTGTCGCCGTC 1824
QY 201 SerIleSerAlaValAlaValLeuAlaPro 210
Db 1825 TCCATTCTGCGGTGCGGTTTGGGCTTTTAGCCCCC 1854
```

Search completed: October 28, 2004, 10:30:32  
Job time : 339 secs

This Page Blank (uspto)